

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS BIOLÓGICAS
Departamento de Genética



TESIS DOCTORAL

**Influencia sobre la composición de ácidos grasos de grasas
subcutánea e intramuscular del SSC12 porcino**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

María Muñoz Muñoz

DIRECTOR:

Luis, dir Silió López

Madrid, 2015



UNIVERSIDAD COMPLUTENSE



5329971422

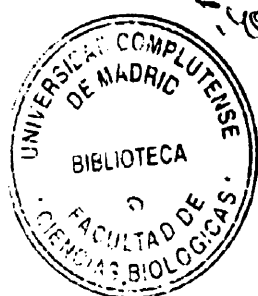


Instituto Nacional de Investigación
y Tecnología Agraria y Alimentaria

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE BIOLOGÍA
DEPARTAMENTO DE GENÉTICA

INFLUENCIA SOBRE LA COMPOSICIÓN DE
ÁCIDOS GRASOS DE GRASAS SUBCUTÁNEA E
INTRAMUSCULAR DEL SSC12 PORCINO

*INFLUENCE OF PORCINE SSC12 IN FATTY ACID
COMPOSITION OF SUBCUTANEOUS AND
INTRAMUSCULAR FAT*



140480867
María Muñoz Muñoz

Madrid, Febrero 2011

FACULTAD DE BIOLOGÍA

A mis padres,
a mi hermana Rosa

En primer lugar, me gustaría empezar agradeciendo a mis codirectores de tesis Luis y Estefânia la confianza depositada en mí para desarrollar este trabajo. Mil gracias por todo lo que me habéis enseñado estos años, por las oportunidades que me habéis brindado de viajar a tantos sitios para realizar estancias, participar en congresos y asistir a cursos que han contribuido a mi formación. Gracias por vuestro trabajo, vuestra dedicación, por preocuparos por mí, por ser algo más para mí que mis “jefes”.

Gracias a todos los que han sido, son y serán mis compañeros del INIA. A Ana por estar siempre ahí, ayudarme, aportarme ideas, echarme cables, por ser mi amiga. A Carmen Rodríguez, por todo lo que me ha enseñado, lo que me ha ayudado siempre en lo que he necesitado y por todo el cariño que he recibido estos años de ella. A Almudena por estar siempre ahí dispuesta a echar una mano (nunca olvidaremos aquellos archivos infernales de millones de filas). A Cristina por los ánimos dados a lo largo de los años. A Carmen Barragán por ayudarme siempre en todo lo que ha podido, por hacerme ver lo necesario de ser organizado y ordenado dentro del laboratorio (y fuera de él). A Amanda, mi compañera de viajes, ¡¡cuántos congresos, cursos y salidas al matadero hemos compartido!!, por estar ahí para cualquier cosa y ser mi amiga (... a pesar de que me obligues a probar el bacalao). A Carmen Amador por hacerme la vida más fácil enseñándome a manejar nuevos programas informáticos, por hacerme ver que la versión original es mejor que la doblada, por estar siempre ahí para escucharme. A Gloria por su apoyo, cariño, amistad. A Dafne, la pizpireta del grupo, por tener siempre esa sonrisa en la cara que tanta alegría me da. A Chari por su espontaneidad y alegría. A todas las chicas/o del laboratorio, Nines, Rita, Yolanda, Fabián y Patricia, por tener siempre todo listo y preparado, la dilución del individuo que me faltaba, la placa de piro para completar mis genotipados, la PCR cuantitativa que necesito para acabar los análisis, sin duda me habéis hecho la vida más fácil. A Jaime por ser tan sencillo y cariñoso y tener siempre una palabra amable que ofrecer. A Beatriz, por preocuparse siempre por cómo me iba y porque creo que hace falta gente como ella que piensa que es posible cambiar las cosas. A toda la pandilla del desayuno o comida, Fernando, Enrique, Judit, Ángeles, Óscar, Evangelina, por todos los momentos en los que nos reímos, discutimos, en definitiva, nos olvidamos por un rato del trabajo. Al resto de compañeros del departamento Jesús, María, Ane, Luis Alberto, Miguel Ángel, Malena, Silvia, Cristina Meneses...

AGRADECIMIENTOS

Muchas gracias al grupo del IRTA en primer lugar por proporcionarnos los datos del escaneado de los microarrays incluidos en esta tesis y por los momentos que hemos compartido en congresos, cursos y sitios menos agradables como las salidas al matadero. Gracias a Noelia, Raquel, David, Quim, José Luis, Romi, Ángela, María, Luis Varona.

También quiero dar las gracias al grupo de la UAB, en primer lugar, por proporcionar los datos de los chips de genotipado de alta densidad así como por los días que pasé allí al poco de comenzar mi periodo de doctorado en el que me estuvieron enseñando sus métodos de trabajo, con los que también he padecido las tempraneras salidas al matadero y con los que también he compartido momentos más agradables. Gracias a Armand, Miguel, Josep María, Marcel, Anna Mercadé, Alí, Ingrid, Oriol...

I am very grateful to Dr. Bo Thomsen, to give me the chance to go to the Genetics and Biotechnology Department at Aarhus University and to Mathilde Nielsen for help me and teaching me at the lab during my training period at Foulum.

To Emmanuelle, for letting me be part of her project during my short visit at INRA in Jouy en Josas and worry about me, always. To Per, for the good moments in Paris, despite he doesn't admit that Rafa Nadal is the best. Gracias a Jordi por todo lo que me ayudó y se preocupó por mí cuando estuve en París. Thanks a lot for my stay in Paris!

Gracias a mis compañeros de comienzos de la facultad, por iniciar juntos este viaje en el año 2000 y pese a que transcurrieran años sin vernos ahora nos reunimos como si el tiempo no hubiese pasado, gracias a Soni, "Jebi", Jorgito, Ivan, Axel, Alfonso, Paloma. A mi gran panda de amigos, por compartir tantos momentos, viajes, "discusiones amigables", risas, llantos durante todos estos años, gracias a Carliños, Cris T, Beñat T, Cris pequeña, Oihane, Ianire, Patricia, Esti. No me olvido de Silvia, Sarita y Jesús, mis compis de piso, con los que he compartido un montón de buenos y malos momentos durante los años de doctorado, por esos "holas" al llegar a casa, los martes de "House", los zombis, las "esquinitas", los "Simpáticos", los partidos del Mundial... y por supuesto muchas gracias a Jesús por llevarme las secuencias al CIB. Y por supuestísimo, muchas gracias a mi Carmelilla, por nuestros comienzos en el laboratorio de *Drosophila*, por las risas, por los llantos, por estar ahí cada vez que la he necesitado, ¡¡muchísimas gracias!!

AGRADECIMIENTOS

Gracias a Sergio, por todos los buenos momentos que hemos pasado juntos, que los malos ya no importan, por todos los viajes, las carcajadas, los paseos, por confiar siempre en que me iban a conceder aquella “beca del cerdo ibérico”. Muchas gracias también a Juan y a Juani por ser siempre tan cariñosos conmigo y preocuparse por mí.

Gracias a mis abuelos Luisa, Inés, Félix y Antonio, los que de alguna manera u otra han ayudado a que esté aquí y los que espero que me estén viendo por un agujerito y se sientan orgullosos de mí. Gracias a todos mis tíos, Pepe, Tono, Elena, Mari Pili, Luis Mariano, Rosa, Tío Paco, Tío Antonio, Tía Piedad y por supuesto Tita Pili, por estar pendientes de mí y ayudarme siempre que han podido con cualquier cosa. ¡Muchas gracias!

Gracias a mi hermana Rosa, de la que cada día aprendo cosas nuevas, a pesar de que yo soy la mayor, por ser tan responsable y por ser tan buena persona, ¡¡hacen falta muchas como tú pequeña!!

Muchas, muchísimas gracias a mis padres, que siempre me han apoyado en todo, aunque muchas veces no estuvieran de acuerdo conmigo o no entendiesen por qué hacía las cosas. Muchas gracias por estar siempre ahí, en los buenos y en los malos momentos. Muchas gracias por educarme de la manera que lo habéis hecho. Papá, gracias por hacer que me esforzase al máximo cada día. Mamá, gracias por ser siempre tan positiva y creer que todo se puede conseguir. Muchísimas gracias a los dos, os quiero mucho.

INDICE

RESUMEN	1
SUMMARY	4
1. INTRODUCCIÓN	7
1.1. LA PARADOJA DE LAS GRASAS ANIMALES	7
1.2. MODIFICACIÓN DE LA COMPOSICIÓN DE LA GRASA A TRAVÉS DE LA ALIMENTACIÓN	9
1.3. MODIFICACIÓN DE LA COMPOSICIÓN DE LA GRASA A TRAVÉS DE LA GENÉTICA	10
1.4. ANTECEDENTES DIRECTOS DEL PRESENTE ESTUDIO	12
1.5. GENÓMICA EN ANIMALES DOMÉSTICOS	14
2. OBJETIVOS	17
3. MATERIALES Y MÉTODOS	18
3.1. MATERIAL ANIMAL	18
3.2. EXTRACCIÓN DE ÁCIDOS NUCLEICOS	25
3.3. ANÁLISIS MOLECULAR DE GENES CANDIDATOS	28
3.4. PCR CUANTITATIVA A TIEMPO REAL	34
3.5. GENOTIPADO A GRAN ESCALA	38
3.6. ANÁLISIS ESTADÍSTICOS	41
3.7. ANÁLISIS DE EXPRESIÓN CON MICROARRAYS	48
RESULTADOS	
4. <i>Caracterización de genes candidatos</i>	54

4.1. Gen <i>Acyl-coA oxidase 1 (ACOX1)</i>	54
4.2. Gen <i>ATP citrate lyase (ACLY)</i>	59
4.3. Gen <i>Sterol regulatory element binding factor 1 (SREBF1)</i>	64
4.4. Gen <i>Acetyl-CoenzymeA carboxylase alpha (ACACA)</i>	70
5. <i>QTLs y genes candidatos para composición de ácidos grasos en grasa subcutánea</i>	
5.1 MAPEO DE LIGAMIENTO	73
5.2 ANÁLISIS DE DETECCIÓN DE QTLs	74
5.3 ANÁLISIS DE ASOCIACIÓN DE SNPs EN GENES CANDIDATOS <i>ACOX1, ACLY y ACACA</i>	77
5.4. OTROS ANÁLISIS DE ASOCIACIÓN	82
5.5. ANÁLISIS DEL POLIMORFISMO <i>ACACA</i> c.5634T>C EN CERDOS IBÉRICOS PUROS Y CRUZADOS CON DUROC	83
5.6. ANÁLISIS DE EXPRESIÓN DEL GEN <i>ACACA</i>	85
6. <i>QTLs y genes candidatos para composición de ácidos grasos en grasa intramuscular</i>	89
6.1. ANÁLISIS ESTADÍSTICOS EN GRASA INTRAMUSCULAR	89
6.2. NUEVOS MAPAS DE LIGAMIENTO DEL CROMOSOMA 12	89
6.3. ANÁLISIS DE DETECCIÓN DE QTLs	92
6.4. ANÁLISIS DE ASOCIACIÓN	101
6.5. ANÁLISIS DEL POLIMORFISMO <i>ACACA</i> : c.5634T>C EN CERDOS IBÉRICOS PUROS Y CRUZADOS CON DUROC	111
7. <i>Análisis de expresión</i>	113

7.1. ANÁLISIS ESTADÍSTICO DE DATOS DE EXPRESIÓN	113
8. DISCUSIÓN	122
8.1. GENES CANDIDATOS	122
8.2. DETECCIÓN CLÁSICA DE QTLs PARA COMPOSICIÓN DE ÁCIDOS GRASOS PARA GRASA SUBCUTÁNEA	128
8.3. DETECCIÓN CLÁSICA DE QTLs PARA COMPOSICIÓN DE ÁCIDOS GRASOS PARA GRASA INTRAMUSCULAR	130
8.4. ANÁLISIS DE ASOCIACIÓN BASADOS EN EL <i>chip</i> <i>SNP60</i>	138
8.5. ANÁLISIS DE EXPRESIÓN	142
CONCLUSIONES	146
CONCLUSIONS	149
BIBLIOGRAFÍA	150
ANEJO	161

RESUMEN

La selección llevada a cabo en porcino durante los últimos años ha favorecido el crecimiento magro a costa de reducir el contenido de grasa subcutánea e intramuscular. Este hecho ha provocado una disminución de la calidad de la carne debida a la pérdida de sus cualidades sensoriales y tecnológicas. Asimismo los cambios en los porcentajes de ácidos grasos saturados (SFA), monoinsaturados (MUFA) y poliinsaturados (PUFA) influyen en las características organolépticas de la carne de cerdo y de sus productos curados. Por otra parte, la composición de ácidos grasos de las grasas animales afecta la salud humana ya que dietas ricas en SFA incrementan el colesterol en plasma y por tanto el riesgo de enfermedades cardiovasculares. En diversos estudios se ha demostrado que la composición de ácidos grasos puede alterarse mediante la alimentación y mediante técnicas genéticas.

El presente trabajo forma parte del consorcio IBMAP, en el que grupos del INIA, IRTA y Universidad Autónoma de Barcelona (UAB) vienen cooperando desde 1996 en proyectos sucesivos con el objetivo de estudiar la base genética de caracteres de crecimiento, conformación, deposición de grasa y metabolismo lipídico. Los objetivos de este estudio son consecuencia de los resultados obtenidos en un trabajo previo en el SSC12 (Muñoz *et al.*, 2007) en el que se identificaron dos QTLs con efecto sobre diferentes ácidos grasos medidos en grasa subcutánea. En este cromosoma se localizan varios genes relacionados con el metabolismo de ácidos grasos, entre ellos el gen *Acetyl CoA carboxylase* (*ACACA*) que es un sólido gen candidato para la mutación causal subyacente a uno de los QTLs.

En este estudio se ha tratado de aumentar el conocimiento sobre la implicación de genes localizados en el SSC12 en la composición de ácidos grasos de las grasas subcutánea e intramuscular. El material biológico utilizado incluye individuos de las generaciones F2, F3, RC1 (♂ F2 x ♀ Landrace) y RC2 (♂ F1 x ♀ Landrace) pertenecientes al cruce experimental Ibérico x Landrace (IBMAP).

Se emplearon diversas estrategias. La primera de ellas consistió en la caracterización, detección y genotipado de polimorfismos en nuevos genes candidatos (*ACOX1*, *ACLY* y *SREBF1*) y en una posible región reguladora del gen *ACACA*. Se detectaron cinco polimorfismos en el gen *ACOX1*, dos en el gen *ACLY* y siete en el gen *SREBF1*. Se construyó un nuevo mapa de ligamiento del cromosoma 12 combinando la información de seis marcadores microsatélite y cuatro SNPs con los haplotipos detectados en el gen *ACOX1*, H1: c.[481G; 1238A; *21G] y H2: c.[481C; 1238G; *21A] así como los SNPs *ACLY*: c.*523T>C y *SREBF1*: c.1023T>C. El mapa obtenido se utilizó en un nuevo

análisis de detección de QTLs que permitió identificar dos regiones cromosómicas con efectos sobre el contenido de diversos ácidos grasos presentes en la grasa dorsal. El primer QTL afecta a los porcentajes de los ácidos linolénico [C18:3, (n-3)], palmitoleico [C16:1(n-9)], palmítico (C16:0) y la Longitud Media de la Cadena (ACL) y el segundo QTL los porcentajes de palmitoleico [C16:1(n-9)], esteárico (C18:0) y vaccénico [C18:1(n-7)]. En los análisis de asociación posteriores se confirmó la asociación significativa entre el SNP *ACACA*: c.5634T>C y los porcentajes de ácidos grasos afectados por el segundo QTL. No se detectaron efectos significativos de los polimorfismos de los genes *ACOX1*, *ACLY* y *SREBF1*. Asimismo, en una región intrónica del gen *ACACA* con homología al Promotor I bovino, se ha identificado un polimorfismo (*ACACA*: g.48001A>G) con efecto sobre el porcentaje de ácido esteárico (C18:0) de la grasa dorsal. Los genotipos de ambos polimorfismos afectan la expresión del gen *ACACA*. La búsqueda de la mutación causal subyacente a este QTL deberá orientarse a los otros dos promotores y a las regiones 5' y 3'UTR cuya secuenciación no se ha realizado todavía en este material.

Se construyeron tres nuevos mapas de ligamiento del SSC12 a partir de la información de algunos de los marcadores genéticos citados y de la proporcionada por el *PorcineSNP60 Genotyping Bead Chip*. En el primero se combinó la información de seis microsatélites, el SNP *SREBF1*: c.1023T>C y 26 SNPs seleccionados del *chip PorcineSNP60*. En el segundo, se utilizó la información de 50 SNPs del *chip PorcineSNP60* y en el tercero los genotipos de 489 SNPs. Estos mapas se emplearon para realizar tres análisis clásicos de detección de QTLs para composición lipídica de la grasa intramuscular, en cerdos de las generaciones F3, RC1 y RC2 del cruce IIBM. Se ha detectado un QTL con efecto sobre el porcentaje de ácido palmítico (C16:0) en la grasa intramuscular. Este resultado se ha obtenido consistentemente en los sucesivos análisis realizados con un número creciente de marcadores (hasta 489 SNPs). El efecto detectado es de una magnitud del 0.6 % por cada alelo *Q* Ibérico, equivalentes a un 2.6% de la media del carácter.

Se comparó la información de los mapas de ligamiento con la de los mapas físicos para estimar la relación entre las distancias genéticas (cM) y físicas (Mb), obteniéndose valores de hasta 3.1 cM/Mb, que exceden la equivalencia 1 cM/Mb convencionalmente aceptada, tanto en el genoma humano como en los de los animales domésticos.

En los análisis de asociación realizados con los genotipos de 978 SNPs con $MAF > 0.15$, se han detectado efectos significativos sobre el porcentaje de ácido palmítico (C16:0) en grasa intramuscular de las sondas *ALGA0066302*, *ASGA0054039*, *ASGA0054041* y *ASGA0054044*. La primera sonda está localizada en el gen *ACACA* y

las otras en el gen *Phosphatidylcholine Transfer Protein (PCTP)*. Las funciones de ambos genes están relacionadas con el metabolismo y transporte de ácidos grasos y el *PCTP* mapea en la posición 29.65 Mb, próxima al intervalo de confianza del QTL (30.96 - 35.02 Mb). Estos resultados sugieren que el *PCTP* es un buen gen candidato funcional y posicional. Por otra parte se detectó un efecto significativo del SNP *H3GA0034898*, localizado en el gen *Arachidonate 12-lipoxygenase (ALOX12)*, sobre el índice C20:4/C18:2. Este gen codifica para la proteína araquidonato 12-lipoxigenasa (12-LOX) que participa en esta ruta de conversión de C20:4 a C18:2, por lo que puede considerarse como un potente gen candidato funcional.

El efecto del genotipo del QTL para el contenido en C16:0 sobre la expresión génica en el músculo *longissimus* ha sido investigado en cerdos de la generación RC2 utilizando el *array* de expresión de *Affimetrix*. Los principales resultados mostraron diferencias de expresión en dos genes del SSC12, *PCTP* y *TMEM97*, con ratios de expresión entre los genotipos *Qq* y *qq* de 0.77 y 1.26, respectivamente. Ambos genes están involucrados en el metabolismo lipídico y mapean dentro del intervalo de confianza del QTL. Sin embargo no se detectaron efectos relevantes sobre la expresión génica del polimorfismo del gen *ACACA* genotipado con la sonda *ALGA0066302*. Los SNPs *ASGA0054039T>C* y *ASGA0054041A>G* del gen *PCTP* mostraron efectos altamente significativos sobre la expresión del propio gen lo que, junto a otros resultados previos, hace que el gen *PCTP* deba ser considerado el principal candidato posicional y funcional subyacente al QTL para el contenido de palmítico en grasa intramuscular.

SUMMARY

SUMMARY

In the last years, the selection performed in pig breeding schemes increased lean growth at the expense of backfat and intramuscular content. Fat content and its fatty acid composition influence the sensorial and technological aspects of meat quality. Changes in the percentages of saturated (SFA), monounsaturated (MUFA) and polyunsaturated (PUFA) fatty acids influence the organoleptic quality of pig meat and dry cured products. Moreover, fatty acid composition of animal fats affects human health since diets with high cholesterol content increase the level of plasmatic cholesterol and therefore the risk of cardiovascular diseases. Several studies reported successful modifications of fatty acid composition of animal fats by feeding changes and genetics factor.

The current work is part of the IBMAP consortium that has cooperated since 1996 in consecutive projects with the goal to study the genetic basis of growth, conformation, fat deposition and lipidic metabolism traits. The objectives of this study are consequence of the obtained results from a previous work performed in SSC12 (Muñoz *et al.*, 2007), in which two QTLs with effect on different backfat fatty acids were identified. Some genes related with fatty acid metabolism are identified in SSC12 one of them was *Acetyl CoA carboxylase (ACACA)*, that it is a strong candidate gene for the causal mutation underlying one of the QTLs.

In the current study, we have tried to increase the knowledge about the implication of SSC12 genes in fatty acid composition of subcutaneous and intramuscular fat. Biological samples from F2, F3, RC1 (♂ F2 x ♀ Landrace) and RC2 (♂ F1 x ♀ Landrace) generation individuals, from an experimental cross Iberian x Landrace (IBMAP) were used.

Different strategies were applied. First of them consists in the characterization, detection and genotyping of polymorphisms located in new candidate genes (*ACOX1*, *ACLY* y *SREBF1*) and in a possible *ACACA* regulatory region. Five polymorphisms were detected in *ACOX1* gene, two in *ACLY* and seven in *SREBF1*. A new SSC12 linkage map was built joining the information from six microsatellites and four SNPs with those from *ACOX1* H1: c.[481G; 1238A; *21G] and H2: c.[481C; 1238G; *21A] detected haplotypes and *ACLY*: c.*523T>C and *SREBF1*: c.1023T>C SNPs. The obtained map was used for a new QTL detection analysis that allowed to identify two chromosomal regions with effects on different backfat fatty acids. The first QTL affects the percents of linoleic [C18:3, (n-3)], palmitoleic [C16:1(n-9)] and palmitic (C16:0) fatty acids and Average Chain Length and the second one QTL the percents of palmitoleic [C16:1(n-9)], stearic (C18:0) and vaccenic [C18:1(n-7)] fatty acids. Also, the significant association between the SNP *ACACA*: c.5634T>C and the fatty acids percents underlying the second QTL was confirmed in the

SUMMARY

subsequent association analysis. Non significant effects of the *ACOX1*, *ACLY* and *SREBF1* polymorphisms were detected. Moreover, a new polymorphism (g.48001A>G) with a significant effect on backfat content of stearic fatty acid (C18:0) has been detected within an intronic region of *ACACA* gene homologous with bovine Promoter I. Both polymorphism genotypes affect the *ACACA* gene expression. The search of the causal mutation underlying the QTL should be focused on the two remaining promoters and the 5' and 3'UTR regions that were not yet sequenced in this material.

Three new SSC12 genetic maps were built from the information of some of the quoted genetic markers and the SNP genotypes supplied by *PorcineSNP60 Genotyping Bead Chip*. The first one combined six microsatellites, the SNP *SREBF1*: c.1023T>C and 26 SNP selected from those genotyped by *PorcineSNP60*. The information of 50 and 489 SNPs was used in the construction of other two maps. These maps were used to perform three classical analyses in order to detect QTL related to fatty acid composition of intramuscular fat on pigs from F3, RC1 and RC2 generations of IBMAP intercross. A QTL effect on palmitic fatty acid (C16:0) percentage in intramuscular fat has been consistently detected in the successive analyses conducted while increasing the number of markers. The detected effect magnitude on the trait was 0.6% for the *Q* Iberian allele, equivalent to 2.6% of the trait mean.

Moreover, two estimations of the equivalence between genetic distances and the Megabase number using the information from the *PorcineSNP60 Genotyping Bead Chips* and the linkage maps were conducted, ranging values of up to 3.1 cM/Mb that exceed the conventionally 1cM/Mb assumed for the human genome as well as in domestic animals.

Significant effects on the palmitic fatty acid (C16:0) percentage in intramuscular fat were detected for the probes *ALGA0066302*, *ASGA0054039*, *ASGA0054041* and *ASGA0054044* in the association analysis performed with the genotypes of 978 SNPs presenting MAF>0.15. The first probe is located in the *ACACA* gene and the others in the *Phosphatidylcholine Transfer Protein (PCTP)* gene. The functions of both genes are related to fatty acid metabolism or lipid transport and *PCTP* maps close to the confidence interval (30.93-35.02 Mb) of QTL, in the 26.95 Mb position. These results suggest that *PCTP* is a good functional and positional candidate gene. In addition, the association analysis results showed a significant effect of the *H3GA0034898* probe on the arachidonic and linoleic ratio (C20:4/C18:2). This probe is located within the *Arachidonate 12-lipoxygenase (ALOX12)* gene. This gene codes arachidonate 12-lipoxygenase (12-LOX) protein that is involved in

SUMMARY

C20:4 to C18:2 conversion pathway. This result provides evidences of its relevance as candidate functional gene.

The effect on the *longissimus* muscle gene expression of the QTL genotype for the palmitic fatty acid content has been investigated in RC2 pigs using the Affymetrix gene expression microarrays. The main results showed expression differences in two SSC12 genes, *PCTP* and *TMEM97*, with expression ratios between *Qq* and *qq* genotypes of 0.77 and 1.26, respectively. Both genes are involved in the lipid metabolism and map within the QTL confidence interval. Non relevant effects of the *ACACA* SNP *ASGA0066302G>A* on gene expression were detected. The SNPs *ASGA0054039T>C* and *ASGA0054041A>G* of the *PCTP* gene showed highly significant effects over its own gene expression, therefore *PTCP* should be considered the main positional and functional candidate gene underlying the QTL for palmitic fatty acid in intramuscular fat.

INTRODUCCIÓN

1.1. LA PARADOJA DE LAS GRASAS ANIMALES

La percepción que los consumidores de los países occidentales tienen de los productos de origen animal integra varios elementos contradictorios. El mundo desarrollado se encuentra actualmente mediatizado por un canon magro en salud y estética, que pretende reducir drásticamente nuestra ingesta de grasa -particularmente la procedente de alimentos de origen animal- favoreciendo el consumo de productos *light*, bajos en calorías, grasas saturadas y colesterol. La etiqueta de insano afecta a alimentos básicos como huevos, lácteos y productos cárnicos. En el caso de estos últimos, ello ha influido, junto a factores como la mejora del índice de transformación del pienso, para que la producción ganadera y la industria hayan rectificado antiguas prácticas, y se orienten desde hace años a obtener carne y productos transformados con un mínimo de grasa visible.

Pero además de los aspectos de los alimentos relacionados con la salud, los consumidores valoran también sus características organolépticas, y son bien conocidas las quejas sobre la pérdida de calidad sensorial de la carne de pollo o de cerdo. Esta disminución de la calidad de carne es resultado del largo proceso de selección de estas especies para aumentar su eficiencia productiva, que ha originado la reducción de su grasa corporal incluida la intramuscular (Lonergan *et al.*, 2001). Está bien establecido que el sabor, jugosidad y ternura de la carne dependen en gran medida de la cantidad y calidad del contenido graso de la misma, determinado por sus principales componentes, los ácidos grasos de cadena larga, que afectan su grado de saturación, estabilidad y aroma. Las características de la grasa, localizada en tejido adiposo y en músculo, son cruciales para la calidad nutricional y sensorial de la carne y sus productos derivados (Webb y O'Neill, 2008).

El consumo excesivo de grasa animal se ha asociado a un aumento del riesgo de tumor colo-rectal, que sin embargo no se ha comprobado en el análisis conjunto de 13 estudios epidemiológicos independientes (Lin *et al.*, 2004). La asociación entre grasas animales y enfermedades cardiovasculares ha sido ampliamente estudiada, lo que ha dado lugar a diferentes recomendaciones médicas desde la exclusión total de las grasas en la dieta a la admisión de un consumo moderado. Pero el énfasis de los consejos dietéticos se ha desplazado cada vez más desde la cantidad de grasa hacia su composición. La grasa animal está compuesta fundamentalmente por ácidos grasos saturados (SFA) y monoinsaturados (MUFA) presentes en los triglicéridos, siendo el más abundante el ácido oleico [C18:1(n-9)] seguido por los ácidos palmítico (C16:0) y esteárico (C18:0) (Valsta *et al.*, 2005). Nótese que los ácidos grasos se caracterizan por la longitud de la cadena de átomos de carbono,

1. INTRODUCCIÓN

generalmente de número par (16, 18, 20 o 22 son los de mayor significación nutricional), y en el caso de los mono y poliinsaturados (PUFA) por la posición del primer doble enlace respecto al metilo terminal (n-3, n-6 o n-9). Los saturados (SFA) con 14 o 16 carbonos (mirístico y palmítico) aumentan el colesterol del plasma, mientras que el de 18 carbonos (esteárico) es neutral y los insaturados (MUFA y PUFA) lo reducen y su consumo se asocia con un menor riesgo de enfermedades cardiovasculares (Lichtenstein, 2003).

Por ejemplo, entre las recomendaciones dietéticas de la *American Health Association* para reducir el riesgo de enfermedades cardiovasculares se incluye que el consumo total de grasa no supere el 30% de la ingesta energética, siendo el de SFA menor del 10% del aporte energético (Krauss *et al.*, 2000). Por el contrario existe evidencia científica de que los ácidos grasos monoinsaturados (MUFA) reducen el riesgo de cardiopatías coronarias (Mente *et al.*, 2009). La *World Health Organization* (2003) recomienda un consumo de SFA inferior al 10% de la energía total de la dieta, de MUFA entre el 10-15% y de PUFA entre el 6 y 7,5% (n-6, 5-8%; n-3, 1-2%) de la energía de la dieta. En cuanto a estos últimos, los nutricionistas subrayan la importancia del equilibrio n-6/ n-3, correspondiente a los PUFA que tienen como respectivos precursores el linoleico [C18:2(n-6)] y el α -linolénico [C18:3(n-3)]. Se aconseja que este cociente debe tener en la dieta humana valores comprendidos entre 4 y 6 (Jimenez-Colmenero *et al.*, 2010).

Aunque pueda ser favorable desde una perspectiva nutricional, un aumento excesivo del contenido en PUFA de los tejidos muscular y graso, resulta indeseable en cuanto a la calidad sensorial y tecnológica de la carne y los productos cárnicos. Según (Lawrence y Fowler, 1997), un alto contenido en ácido linoleico [C18:2(n-6)] de la carne de cerdo es valorado negativamente por los consumidores. La degradación de lípidos, esencialmente por oxidación de ácidos grasos insaturados libres (en especial MUFA) durante el procesado de productos cárnicos porcinos, lleva a la formación de numerosos compuestos aromáticos volátiles, algunos de los cuales contribuyen positivamente al aroma de los productos curados (Ruiz *et al.*, 1999). Por el contrario, un excesivo contenido de PUFA puede dar lugar al enranciamiento de la grasa con olor y color indeseables, así como a una disminución de su consistencia asociada al cociente esteárico/linoleico (Cava *et al.*, 1999).

1. INTRODUCCIÓN

1.2. MODIFICACIÓN DE LA COMPOSICIÓN DE LA GRASA A TRAVÉS DE LA ALIMENTACIÓN

La estrategia más utilizada para producir carne con un perfil de ácidos grasos más saludable es la alteración del contenido en ácidos grasos de la ración de los animales, aumentando el porcentaje de MUFA o los cocientes PUFA/SFA o mejor n-3/ n-6 de la misma. Esta opción ha sido ampliamente ensayada particularmente en aves y cerdos, siendo en principio menos efectiva en los rumiantes en los que los ácidos grasos insaturados de la dieta son ampliamente hidrogenados por los microorganismos del rumen (Webb y O'Neill, 2008). Un ejemplo bien conocido es el sistema tradicional de engorde de cerdos ibéricos en montanera, basado en la ingesta de bellotas con más de 6% de grasa de alto contenido de ácido oleico [C18:1(n-9)] (>60%) y de hierba con alto contenido de α -linolénico [C18:3(n-3)], que se refleja en la composición de la grasa corporal de los animales y sus productos (Lopez-Bote, 1998). Esta característica ha sido la base del sistema oficial de verificación del tipo de engorde de cerdos ibéricos que se ha venido utilizado hasta hace pocos años.

Este sistema tradicional de engorde se ha intentado mimetizar en cerdos confinados utilizando diferentes piensos enriquecidos en grasa insaturada. Inicialmente se utilizaron grasas vegetales con alto contenido de ácido linoleico (girasol, soja, maíz), que originan una grasa aún más insaturada que la obtenida en el sistema de montanera y con propiedades inadecuadas para la curación de los productos: largo tiempo de curación, producción de compuestos aromáticos indeseables como hexanal, pentanal, etc (Lopez *et al.*, 1992). Posteriormente se ensayaron piensos enriquecidos con MUFA, en especial oleico, mediante el uso de subproductos de almazaras (oleínas), manteca de cerdos ibéricos de montanera o girasol de alto contenido en oleico (Lopez-Bote, 1998). Esta última es la opción de enriquecimiento de la dieta más estudiada en cerdos ibéricos y de otras razas, a veces en combinación con el suministro de α -tocoferol (Ventanas *et al.*, 2007). La aplicación de estas prácticas no es siempre reproducible y es también objeto de controversia en cuanto a su incidencia en aspectos de la curación de jamones (Jimenez-Colmenero *et al.*, 2010).

Es conveniente que los efectos de la dieta sobre la composición de los tejidos animales se analicen conjuntamente con algunos efectos genéticos. Ovilo *et al.* (2010) (2010) compararon el efecto de dietas isocalóricas con diferente porcentaje de girasol de alto contenido en oleico (0 vs 6%) sobre la composición lipídica y la expresión del gen *SCD* ($\Delta 9$ desaturasa) en grasa subcutánea dorsal y en el músculo *longissimus* de cerdos ibéricos. Se detectaron diferencias significativas entre ambos grupos en el perfil de ácidos grasos de la grasa subcutánea pero no en grasa intramuscular. Los resultados del análisis de expresión

1. INTRODUCCIÓN

por PCR cuantitativa mostraron una gran estabilidad de la expresión del gen *SCD* en la grasa subcutánea, y una mayor transcripción en músculo en los cerdos alimentados con la dieta no enriquecida en ácido oleico. Dichos resultados indican que el control genético de la composición lipídica pudiera ser distinto en tejidos magros y grasos.

1.3. MODIFICACIÓN DE LA COMPOSICIÓN DE LA GRASA A TRAVÉS DE LA GENÉTICA

La magnitud del componente genético de la composición de ácidos grasos ha sido analizada en las especies ganaderas más importantes. En cerdos de tipo pesado, es conocido el efecto de la raza o cruce sobre la composición de ácidos grasos de la grasa subcutánea en animales destinados a la producción de jamón de Parma. Los tipos genéticos procedentes de cruces de líneas seleccionadas tienen un mayor contenido de PUFA y menor de SFA que los animales obtenidos en cruces tradicionales Large White x Landrace (lo Fiego *et al.*, 2005). Asimismo, en la producción de animales cruzados Duroc x Ibérico la línea genética de la raza Duroc utilizada tiene una notable influencia sobre los porcentajes de SFA y MUFA en grasa subcutánea (Muñoz *et al.*, 2010).

En cuanto a la variación genética dentro de raza, (Sellier y Monin, 1994) han revisado tres estudios con valores medios de las heredabilidades de 0.58 para el ácido linoleico y 0.51 para el esteárico, que sugieren que el contenido en los ácidos grasos tiene un fuerte componente genético. (Fernández *et al.*, 2003), a partir de un gran volumen de información registrada en cerdos ibéricos por la asociación de ganaderos AECERIBER, han obtenido estimas de heredabilidad del porcentaje en grasa subcutánea de los cuatro ácidos grasos más abundantes (palmítico, esteárico, oleico y linoleico), así como de sus correlaciones genéticas con el porcentaje de grasa intramuscular, el crecimiento diario y el rendimiento de jamones, paletas y lomos. Los valores de la heredabilidad de los ácidos grasos oscilan entre 0.29 y 0.41, siendo la heredabilidad del porcentaje de grasa intramuscular 0.25. Las correlaciones genéticas estimadas indican algún antagonismo genético entre el porcentaje de grasa intramuscular y el rendimiento de jamones y lomos y una correlación positiva no deseable ($r=0.31$ a 0.57) del contenido en ácido linoleico con el rendimiento de las distintas piezas nobles. Estos resultados indican que, al menos a corto plazo, sería posible realizar un esquema de selección efectivo para aumentar el porcentaje de piezas nobles en las canales de cerdo ibérico, que debería priorizar la mejora del rendimiento en paletas ya que su bajo rendimiento es la principal limitación de las canales

1. INTRODUCCIÓN

de ibéricos puros. Sin embargo, a medio plazo, para evitar el posible deterioro de la calidad de carne, el objetivo del esquema de selección debería modificarse para incorporar al objetivo de selección los contenidos de grasa intramuscular (GIM) y de ácido linoleico.

La revolución genómica iniciada hace dos décadas ha permitido iniciar el estudio de la arquitectura genética de los caracteres de interés económico en las especies utilizadas en producción ganadera. El conocimiento de los genes y variantes génicas involucrados en la variación genética de la composición de ácidos grasos será de gran ayuda en el futuro diseño de nuevas tácticas de selección para estos caracteres, y a corto plazo una útil herramienta complementaria de la selección porcina convencional.

Un gran número de loci con efecto sobre caracteres cuantitativos (QTLs) ha sido identificado en los genomas de los principales animales domésticos. La detección de QTLs consiste en la localización de regiones genómicas que contienen uno o más genes que afectan a la variación de un carácter cuantitativo (Andersson, 2001). Para llevar a cabo esta aproximación se asume que las líneas parentales son homocigotas para alelos alternativos del QTL (*QQ* y *qq*) y la progenie híbrida se compara en función de los alelos recibidos de las líneas parentales en las regiones genómicas de interés. El método se divide en dos pasos: 1) Cálculo de la probabilidad de que un individuo tenga un determinado genotipo para el QTL en cada posición analizada basada en el genotipo de los marcadores genotipados en la población, y 2) Uso de estas probabilidades para estimar la posición del QTL y efectos del mismo. Los análisis normalmente se realizan estimando la presencia o ausencia del presunto QTL a intervalos de un cM, reteniendo la posición más significativa y la estimación del efecto del QTL en ese punto. En cerdos, el PigQTLdb (<http://www.animalgenome.org/cgi-bin/QTLdb/SS/index>) contiene información de 6.344 QTLs de 281 publicaciones científicas y representando 593 caracteres distintos.

Varios de estos QTLs presentan efectos sobre el contenido de ácidos grasos en grasa subcutánea. El primero de ellos se detectó en la generación F2 del cruce experimental Ibérico x Landrace objeto de la presente tesis doctoral. (Perez-Enciso *et al.*, 2000) identificaron un QTL en el cromosoma 4 con un significativo efecto sobre el contenido de ácido linoleico. (Clop *et al.*, 2003) publicaron después los resultados del barrido genómico completo de este material, en el que detectaron además QTLs en los cromosomas 8 (ácidos palmítico y palmitoleico), 10 (ácido mirístico), y 12 (ácido linoleico). Posteriormente estudios realizados por grupos de Corea y Japón en otros cruces experimentales han descrito QTLs con efectos sobre ácidos grasos en grasa subcutánea en los cromosomas 1, 4, 5, 7, 9, 15 y 17 (Kim *et al.*, 2006; Nii *et al.*, 2006) .

1. INTRODUCCIÓN

En los barridos genómicos llevados a cabo con marcadores microsatélite, las regiones cromosómicas donde se localizan por primera vez los QTLs son generalmente grandes, por lo que es aconsejable disponer de una densidad mayor de marcadores y mapas en los que pueda delimitarse su posición con mayor precisión (Fan *et al.*, 2010). Este problema resulta especialmente grave en algunas regiones cromosómicas en las que no se dispone de marcadores microsatélites o estos no son informativos. En este sentido, las nuevas técnicas de secuenciación del genoma porcino y genotipado a gran escala han generado decenas de miles de polimorfismos nucleotídicos (SNPs) distribuidos por todos los cromosomas (SNPs) que facilitan el refinamiento de la localización de los QTLs. Los estudios de QTLs se continúan con la búsqueda y selección de genes candidatos funcionales (función bioquímica conocida y relacionada con el carácter de interés) y/o posicionales (localizados en el intervalo de confianza del QTL) y posteriormente, en la detección de polimorfismos dentro del gen. Los polimorfismos detectados y genotipados permiten realizar análisis de asociación para establecer la relación estadística entre las medidas de los caracteres a estudiar y el genotipo de los mismos. Sin embargo la identificación de las mutaciones causales de los QTLs es un problema complejo en animales domésticos.

Los métodos que han tenido éxito en especies de animales modelo no son aplicables en animales domésticos por diversas causas (largos intervalos generacionales, dificultad de disponer de transgénicos y líneas *knock-out*). Hay que recurrir a otros enfoques menos potentes para obtener evidencia a favor de la causalidad de una mutación en un gen candidato. Los más comunes son la validación genética en distintas poblaciones del efecto atribuido a un polimorfismo, y la validación funcional mediante la comparación de las variantes genéticas en cuanto a su transcripción génica o funcionalidad de la proteína producto. No es sorprendente que solo dos mutaciones causales con efectos sobre caracteres de interés productivo hayan sido validados hasta el momento en el genoma porcino: en los genes *PRKAG3* del cromosoma 15 (calidad de carne) e *IGF2* en el 2 (rendimiento magro).

1.4. ANTECEDENTES DIRECTOS DEL PRESENTE ESTUDIO

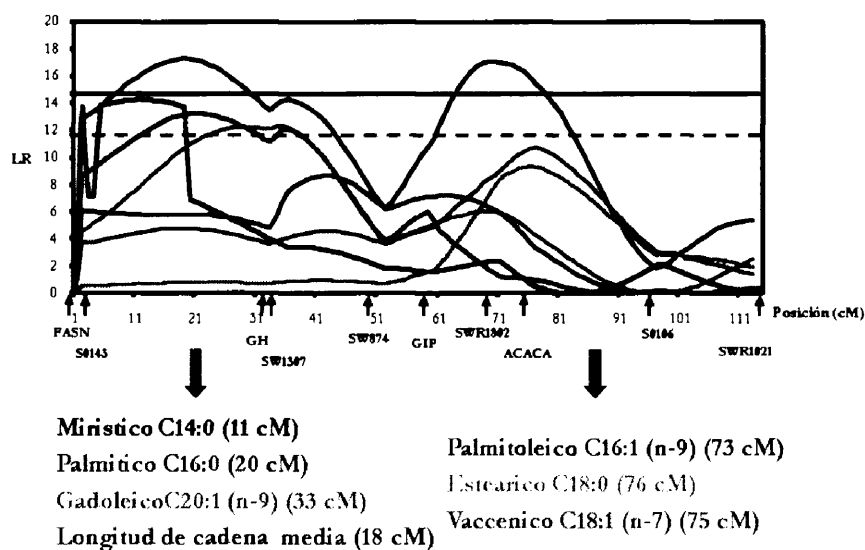
El presente estudio se ha realizado en el marco del proyecto CPE03-010-C3 y del subproyecto AGL2008-04818-C03-03, en colaboración con grupos de la Facultad de Veterinaria de la Universidad Autónoma de Barcelona y del IRTA-Lleida, a cargo de otros

1. INTRODUCCIÓN

dos subproyectos coordinados. Los tres grupos vienen cooperando desde 1996 en proyectos sucesivos en los que se han ido generando distintas generaciones del cruce experimental IBMAP (Ibérico Guadyerbas x Landrace), y se han detectado numerosos QTLs, y asociaciones con mutaciones en distintos genes candidatos subyacentes, para caracteres de crecimiento, conformación, deposición de grasa y metabolismo lipídico. Actualmente se llevan a cabo diversos estudios de asociación y funcionales en genes candidatos biológicos y posicionales relacionados con los QTLs para contenido de ácidos grasos citados en el anterior apartado: genes de la familia *Fatty acid binding proteins* *FABP4* y *FABP5* para el QTL del cromosoma 4, *Microsomal triglyceride transfer protein large subunit* (*MTTP*) como gen candidato para el QTL del cromosoma 8, y *Long chain acylCoA synthetase* 4 (*ACSL4*) seleccionado como gen candidato para un QTL en el cromosoma X, con efecto sobre el porcentaje de grasa intramuscular.

En este contexto se plantearon los objetivos de esta tesis doctoral, a partir de los resultados de un estudio previo orientado específicamente a la detección en el material IBMAP de QTLs en el SSC12 para composición de ácidos grasos de la grasa subcutánea y el análisis de asociación de polimorfismos en genes candidatos (Muñoz *et al.*, 2007). Un nuevo análisis de ligamiento basado en 10 marcadores genéticos permitió identificar dos QTL. El primer QTL con efectos sobre los porcentajes de los ácidos mirístico (C14:0), palmítico (C16:0) y gadoleico [C20:1, n-9]. El segundo QTL afecta a los porcentajes de palmitoleico [C16:1, n-9], esteárico (C18:0) y vaccénico [C18:1, n-7] (Figura 1).

Figura 1. Perfiles de la razón de verosimilitud en el SSC12 obtenidos en el trabajo de Muñoz *et al.* (2007)



1. INTRODUCCIÓN

Estos resultados justificaron el estudio de los siguientes genes como candidatos posicionales y funcionales.

- *Fatty acid synthase (FASN)*. Este gen codifica para una enzima multifuncional que es la principal responsable de la biosíntesis *de novo* de ácidos grasos.
- *Gastric inhibitory polypeptide (GIP)*. Este gen codifica para hormona que pertenece a la superfamilia del glucagon. La función de ésta es mantener la homeostasis de la glucosa como un potente estimulador de la secreción de la insulina jugando un papel principal en la acumulación de grasa en los tejidos adiposos.
- *Acetyl-Coenzyme A carboxylase alpha (ACACA)*. Este gen codifica para una enzima limitante en la biosíntesis de ácidos grasos que cataliza la carboxilación de la acetil-CoA a malonil-CoA.

Pese a haberse identificado tres polimorfismos del gen *FASN* que daban lugar a cambios aminoacídicos, el análisis de este gen como candidato al primer QTL resultó fallido. Los análisis de asociación de posibles mutaciones subyacentes al segundo QTL mostraron efectos significativos de signo opuesto del polimorfismo *GIP*: c.77A>G sobre los ácidos palmitoleico y esteárico. Asimismo se detectaron efectos significativos y de distinto signo del polimorfismo *ACACA*: c.5634T>C sobre el ácido esteárico y sobre los monoinsaturados palmitoleico y vaccénico. En el trabajo actual nos propusimos la profundización de este estudio a partir de la obtención de nuevas generaciones del cruce experimental y la utilización de nuevas técnicas de genotipado masivo que permiten disponer de un mapa cromosómico de alta densidad, y de análisis de expresión génica mediante microarrays, que ofrecen la posibilidad de identificar nuevos genes candidatos.

1.5. GENÓMICA EN ANIMALES DOMÉSTICOS

El desarrollo del proyecto de secuenciación del genoma humano ha impulsado el desarrollo de nuevas técnicas y metodologías genómicas que han sido transferidas a estudios genómicos en animales domésticos (Goddard y Hayes, 2009). Las técnicas de secuenciación masiva permiten la secuenciación de genomas enteros mediante tres pasos: fragmentación del genoma en secuencias cortas, lectura de estas secuencias y unión de las mismas a partir de un genoma de referencia. Actualmente, existe una amplia oferta comercial de plataformas que permiten llevar a cabo esta técnica, pero todas ellas tienen en común esta metodología (Metzker, 2010). Estas técnicas están siendo aplicadas en la

1. INTRODUCCIÓN

finalización de la secuenciación del genoma porcino así como el de otras especies domésticas.

Tabla 1.1. Tabla extraída de (Fan *et al.*, 2010). *Chips* de SNPs de Illumina desarrollados hasta el momento para los animales domésticos más importantes. El intervalo es la distancia media entre SNPs y MAF, la frecuencia mínima del alelo menor media medida en diferentes poblaciones

Especies	Chip	Nº SNPs	Nº SNPs mapeados	Intervalo (kb)	MAF	Estado
Gallina	Diversos*	-	-	-	-	Acceso libre con restricciones
Perro	<i>CanineSNP20</i>	22,362	22.000 (CanFam2.0)	125	0.27	Comercialmente disponible
Perro	<i>CanineHD</i>	170,000	170.000 (CanFam2.0)	14.30	0.23	Comercialmente disponible
Vacuno	<i>BovineSNP50</i>	54,001	52,255 (Btau4.0)	51.50	0.25	En desarrollo
Vacuno	<i>BovineHD</i>	>500,000	-	-	-	En desarrollo
Vacuno	<i>Bovine3K**</i>	3,000	3.000 (Btau4.0)	-	0.21	Comercialmente disponible
Caballo	<i>EquineSNP50</i>	54,602	54.602 (EquCab2.0)	43.20	-	Comercialmente disponible
Cerdo	<i>PorcineSNP60</i>	64,232	55.446 (Sscrofa9)	40.70	0.27	Comercialmente disponible
Oveja	<i>OvineSNP50</i>	54,241	-	46	0.30	Comercialmente disponible

* Diversos *chips* contruidos, incluyendo uno de 60K.

** SNPs seleccionados del *chip BovineSNP50* enfocados a la potencial selección de vacuno de leche

A partir de los datos generados de la secuenciación de los genomas completos llevados a cabo hasta el momento se han desarrollado plataformas o *chips* de SNP de alta densidad en diferentes especies domésticas. En la tabla 1.1 se muestra una recopilación de estos *chips*. Los millares de genotipos generados con esta técnica permiten abordar los estudios genéticos en animales domésticos con un volumen de información genotípica impensable hasta el momento. La aplicación más popular de estas técnicas de genotipado masivo es la denominada Selección Genómica que ofrece nuevas expectativas a los sistemas de selección convencional (Meuwissen *et al.*, 2001). En esta tesis hemos utilizado los cientos de genotipos obtenidos con una de estas plataformas a la obtención de un nuevo mapa de ligamiento del cromosoma 12 porcino, a la detección de QTLs para ácidos

1. INTRODUCCIÓN

grasos en grasa intramuscular con un enfoque clásico y asimismo a la realización de un estudio de asociación a nivel cromosómico.

Los estudios de asociación genómica (GWAS) se han aplicado en humanos, para la detección de factores genéticos relacionados con ciertas enfermedades. En especies domésticas, no son muchos los estudios de asociación genómica publicados hasta la fecha (Duijvesteijn *et al.*, 2010; Gorbach *et al.*, 2010), debido probablemente a la reciente salida al mercado de los *chips* de alta densidad. Sin embargo, en los últimos congresos internacionales celebrados entre especialistas de genética y mejora de animales domésticos, como la 32 Conferencia de la *International Society for Animal Genetics* (Edimburgo, 26-30 Julio 2010) o el 9º Congreso Mundial de Genética Aplicada a la Producción Animal (Leipzig, 1-6 Agosto de 2010) se han presentado diversos estudios de asociación a nivel genómico en las especies domésticas de mayor importancia económica.

OBJETIVOS

2. OBJETIVOS

El objetivo general de la tesis consiste en una profundización de los factores genéticos localizados en el cromosoma 12 porcino, implicados en la variación de la composición de ácidos grasos de las grasas subcutánea e intramuscular. Para llevar a cabo este objetivo general se acometieron los siguientes objetivos parciales:

1. Confirmación de los QTLs para la composición de ácidos grasos de la grasa dorsal previamente detectados en este cromosoma en la generación F₂ de un cruce experimental Ibérico x Landrace (IBMAP), mediante la realización de un nuevo análisis incluyendo nuevos marcadores.
2. Caracterización del ADNc de los genes candidatos *ACOX1*, *ACLY* y *SREBF1* con funciones relacionadas con el metabolismo de ácidos grasos que puedan explicar los efectos genéticos citados sobre la composición de la grasa dorsal. Detección y genotipado de polimorfismos en los citados genes y en regiones reguladoras del gen *ACACA* explicativos de los mismos efectos. Asociación con el contenido de ácidos grasos en grasa subcutánea y validación de los resultados significativos mediante estudios en otras poblaciones.
3. Detección de QTLs con efecto sobre la composición de ácidos grasos de la grasa intramuscular, en las generaciones F3 y retrocruces del IbmAP, mediante una aproximación clásica utilizando genotipos obtenidos con la plataforma *PorcineSNP60 Genotyping BeadChip*.
4. Realización de un estudio de asociación a nivel cromosómico para los mismos caracteres utilizando los anteriores datos de genotipado de SNPs localizados en el cromosoma 12.
5. Estudio de los efectos sobre la expresión génica en músculo *longissimus* de los SNPs significativamente asociados con el contenidos de ácidos grasos en grasa intramuscular.

MATERIALES Y MÉTODOS

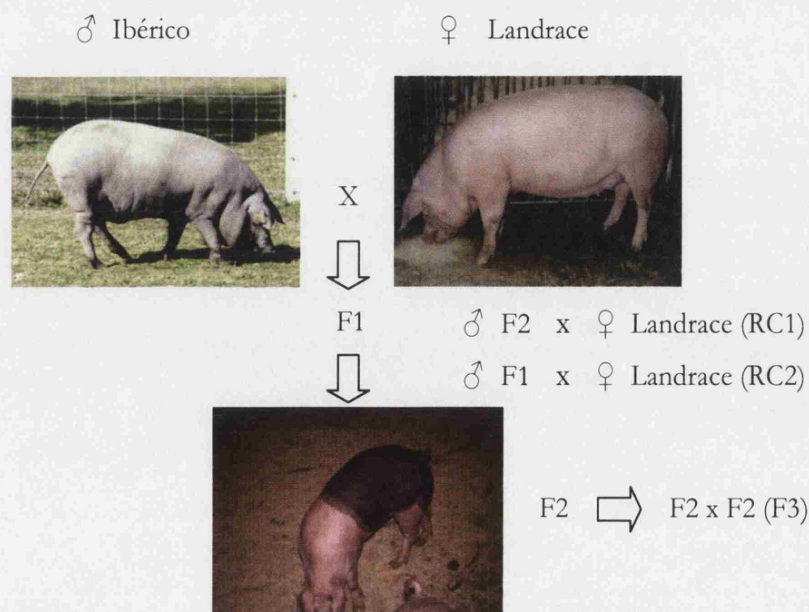
3. MATERIALES Y MÉTODOS

3.1. MATERIAL ANIMAL

3.1.1. Cruce Ibérico x Landrace

El cruce experimental Ibérico x Landrace (IBMAP), fue desarrollado a partir de la línea Guadyervas de cerdos ibéricos (Dehesón del Encinar, Toledo) y de una línea Landrace. La línea Guadyervas representa al tipo negro lampiño de la Vega del Guadiana y se caracteriza por tener una composición corporal extremadamente grasa y elevada consanguinidad (Toro *et al.*, 2000). La línea Landrace usada es una línea maternal, magra, no consanguínea y procede de la granja experimental Nova Genética SA (Solsona, Lleida). Estas dos líneas son extremadamente divergentes para caracteres relacionados con la calidad de carne como grasa intramuscular, espesor de grasa dorsal y crecimiento (Serra *et al.*, 1998) haciendo que el cruce entre ambas sea un excelente material para el estudio de la base genética de caracteres de calidad de carne. A lo largo de los años sucesivos se obtuvieron distintos cruzamientos (Figura 3.1) entre estas líneas y sus progenies (F1, F2, F3, RC1 y RC2).

Figura 3.1. Cruce Ibérico x Landrace



3. MATERIALES Y MÉTODOS

3.1.1.1. Generación F2

Inicialmente, tres machos Ibéricos de la línea Guadyerbas fueron cruzados con 30 reproductoras de la raza Landrace. Posteriormente, seis machos y 64 hembras de la generación F1 fueron cruzados para producir la generación F2. La obtención de la línea, rasgos del pedigrí, manejo de los animales y sacrificio de los mismos fueron descritos en Ovilo *et al.* (2000) y Pérez-Enciso *et al.* (2000). Los animales F2 con registros de composición de ácidos grasos en grasa subcutánea fueron 403. En la tabla 3.1 se muestra un resumen de los principales estadísticos de los caracteres analizados en el presente trabajo.

Tabla 3.1. Estadísticos de los caracteres analizados en la generación F2. Los porcentajes de ácidos grasos corresponden a grasa subcutánea.

Carácter	Media	Desviación típica	Mínimo	Máximo
Peso canal, Kg	74.69	10.21	41.80	107.10
GRIN, %	1.43	0.62	0.36	4.15
C14:0, %	1.51	0.17	0.83	2.12
C16:0, %	21.04	1.41	16.61	25.22
C16:1 (n-9), %	2.51	0.40	1.54	4.07
C18:0, %	10.93	1.29	7.85	25.41
C18:1 (n-9), %	44.42	1.83	37.80	50.33
C18:1 (n-7), %	3.01	0.37	2.40	5.84
C18:2 (n-6), %	14.03	1.66	9.37	18.79
C18:3 (n-3), %	1.08	0.18	0.10	2.23
C20:1 (n-9), %	0.85	0.20	0.22	2.29
C20:2 (n-6), %	0.62	0.16	0.12	0.19
ACL	17.50	0.04	17.40	17.70
DBI	0.83	0.03	0.68	0.93
UI	2.50	0.24	1.50	3.25
PI	18.07	1.89	12.75	23.60

Donde GRIN es el porcentaje de grasa intramuscular determinada por espectroscopia de infrarrojo cercano (NIRS) en una muestra de músculo *longissimus*

3. MATERIALES Y MÉTODOS

lumborum. En una muestra de grasa dorsal tomada a la altura de la 3-4 costilla se determinó la composición de ácidos grasos mediante cromatografía de gases. Los ácidos grasos medidos fueron: mirístico (C14:0), palmítico (C16:0), palmitoleico [C16:1 (n-9)], esteárico (C18:0), oleico [C18:1 (n-9)], vaccénico [C18:1 (n-7)], linoleico [C18:2 (n-6)], linolénico [C18:3 (n-3)], gadoleico [C20:1 (n-9)] y eicosadienoico [C20:2 (n-6)]. Además, cuatro índices metabólicos fueron calculados como se describe en Pamplona *et al.* (1998): longitud media de la cadena (ACL), índice de dobles enlaces (DBI), índice de insaturación (UI) e índice de peroxidabilidad (PI).

3.1.1.2. Otras generaciones del cruce IBMAP: F3, Retrocruce 1 y Retrocruce 2

A partir del apareamiento entre tres verracos y 15 hembras F2 se generaron 56 animales F3 y del apareamiento entre cuatro verracos F2 con 22 hembras Landrace resultó un primer retrocruce (RC1) con 79 individuos con registros disponibles. Los animales de la F2 elegidos como reproductores eran una muestra representativa de la población F2. Posteriormente, un segundo retrocruce (RC2) compuesto por 144 individuos fue generado a partir de la inseminación de 25 hembras Landrace con semen congelado de cinco verracos F1. El engorde de estos animales se realizó en una granja comercial aunque con un manejo similar al de la citada granja experimental Nova Genética SA. Los sacrificios se realizaron en tres lotes en la F3, dos lotes en el RC1 y tres lotes en el RC2, en diferentes mataderos y salas de despiece.

En la tabla 3.2 se muestra un resumen de los registros fenotípicos analizados en estas tres poblaciones. Se realizaron cuatro pesadas desde el nacimiento hasta el sacrificio de los mismos en el caso de las generaciones F3 y RC1 y siete en el caso de la generación RC2. El perfil de ácidos grasos, en este caso, fue analizado mediante cromatografía de gases en una muestra de grasa intramuscular obtenida de músculo *longissimus dorsi*. Se analizaron los mismos ácidos grasos que en la F2 y además se midieron, además del otro isómero del palmitoleico [C16:1 (n-7)] los ácidos margárico (C17:0), heptadecenoico (C17:1), gadoleico (C20:1), eicosatríenoico [C20:3 (n-6)] y araquidónico (C20:4). Por otro lado, además de los índices metabólicos calculados en la F2, se calcularon los índices: porcentaje de ácidos grasos saturados (SFA), porcentaje de ácidos grasos monoinsaturados (MUFA), porcentaje de ácidos grasos poliinsaturados (PUFA), proporción de ácidos grasos monoinsaturados frente a saturados (MUFA/SFA), proporción de ácidos grasos poliinsaturados frente a

3. MATERIALES Y MÉTODOS

saturados (PUFA/SFA) e índice de conversión de ácido linoleico en ácido araquidónico (C20:4/C18:2).

Tabla 3.2. Estadísticos de los caracteres analizados en las generaciones F3, RC1 y RC2. Los porcentajes de ácidos grasos corresponden a grasa intramuscular (GRIN).

Carácter	Media	Desviación típica	Mínimo	Máximo
Peso canal, Kg	74.41	11.10	43.60	105.61
GRIN	1.52	0.80	0.28	4.50
C14:0, %	1.19	0.18	0.76	1.86
C16:0, %	22.66	1.55	18.75	28.32
C16:1 (n-9), %	0.36	0.09	0.06	1.04
C16:1 (n-7), %	2.55	0.54	0.32	4.31
C17:0, %	0.25	0.07	0.12	0.59
C17:1, %	0.46	0.51	0.00	2.43
C18:0, %	13.52	1.38	2.05	17.15
C18:1 (n-9), %	37.29	4.03	28.20	45.78
C18:1 (n-7), %	3.79	0.39	0.34	4.84
C18:2, %	12.27	3.14	6.24	20.54
C18:3 (n-3), %	0.61	0.23	0.22	2.64
C20:0, %	0.23	0.12	0.01	1.42
C20:1, %	0.70	0.19	0.00	1.18
C20:2, %	0.46	0.26	0.00	4.18
C20:3 (n-6), %	0.47	0.39	0.00	5.36
C20:4, %	2.55	1.55	0.00	7.00
ACL	17.40	0.20	16.57	19.72
PI	26.27	9.42	10.37	52.38
DBI	0.84	0.09	0.67	1.09
UI	2.24	0.35	1.56	3.24
SFA	37.84	2.07	29.73	44.16
MUFA	45.11	4.30	34.19	53.97
PUFA	16.39	4.79	7.81	28.72
MUFA/SFA	1.19	0.11	0.92	1.77
PUFA/SFA	0.44	0.15	0.19	0.86
C20:4/C18:2	0.20	0.08	0.01	0.37

3.1.2. Poblaciones comerciales

Por último se ha dispuesto de información generada en un núcleo de selección de cerdas ibéricas puras propiedad de una empresa comercial. Previamente se descartaron aquellas reproductoras ibéricas portadoras de alelos de los genes *IGF2* y *MC1R*, indicadores de introgresión de genotipos foráneos. Se analizaron dos ensayos.

3.1.2.1. Ensayo 1: *Ibérico × Ibérico*: A partir del apareamiento entre 11 verracos y 63 hembras se generaron 174 animales Ibéricos puros. La genealogía de estos animales fue controlada y se identificaron los animales al destete alrededor de los 28 días de edad, mediante la implantación de dispositivos electrónicos inyectables colocados en la cavidad abdominal. Los animales fueron trasladados y manejados hasta su sacrificio en el Centro de Pruebas de Porcino del ITACyL (Hontalbilla, Segovia), donde se controlaron sus rendimientos productivos (crecimiento, consumo, índice de conversión, espesor de tocino dorsal y área de lomo) hasta el sacrificio. Los sacrificios fueron realizados en cuatro lotes. La toma de datos de estudio de la canal y perfil de ácidos grasos fue realizada en la Estación Tecnológica de la Carne de Guijuelo del ITACyL (Guijuelo, Salamanca). Se midió el peso de las canales que tuvo una media de 112.86 ± 10.81 Kg. En la tabla 3.3 se resumen los registros fenotípicos analizados. El perfil de ácidos grasos fue analizado mediante cromatografía de gases en una muestra de 300 g de músculo *longissimus dorsi* cogido a la altura de la última costilla de la canal en el total de los individuos y en la grasa subcutánea de la parte posterior de este músculo, en un grupo de 108. Nótese que se analizó un mayor número de ácidos grasos en grasa intramuscular que en subcutánea. Además de varios ácidos grasos analizados en el cruce IBMAP se midieron los siguientes ácidos: ácido cáprico (C10:0), ácido butírico (C12:0), ácido docosahexaenoico [C22:4 (n-6)] y ácido docosapentaenoico [C22:5 (n-3)].

3.1.2.1. Ensayo 2: *Duroc × Ibérico*: Ocho verracos de la raza Duroc fueron elegidos de un total de 18 machos en base a la heterocigosidad de sus genotipos para los genes *MC4R*, *LEPR* e *IGF2*. Estos verracos se cruzaron con 31 hembras Ibéricas puras para producir un total de 117 cerdos. La genealogía de estos animales fue controlada y se identificaron los animales al destete alrededor de los 28 días de edad, mediante la implantación de dispositivos electrónicos inyectables colocados en la cavidad abdominal. Al igual que en la población 1, los animales fueron manejados hasta su sacrificio en el Centro

3. MATERIALES Y MÉTODOS

de Pruebas de Porcino del ITACyL (Hontalbilla, Segovia) controlándose sus rendimientos productivos (crecimiento, consumo, índice de conversión, espesor de tocino dorsal y área de lomo) hasta el sacrificio.

Tabla 3.3. Perfil de ácidos grasos analizado en grasa subcutánea e intramuscular en animales ibéricos puros del ensayo 1.

Carácter	Grasa subcutánea		Grasa intramuscular	
	Media	Desv. típica	Media	Desv. típica
C10:0	-	-	0.11	0.02
C12:0	0.07	0.01	0.09	0.09
C14:0	1.32	0.14	1.48	0.15
C16:0	22.93	1.47	25.82	2.09
C16:1	2.39	0.36	4.32	0.61
C17:0	0.30	0.08	0.13	0.05
C17:1	0.29	0.09	0.17	0.06
C18:0	11.17	1.28	11.27	1.30
C18:1 (n-9)	50.95	1.89	46.66	2.24
C18:1 (n-7)	-	-	4.47	0.53
C18:2 (n-6)	8.30	0.53	3.59	0.60
C18:3 (n-3)	0.51	0.04	0.19	0.04
C18:2 (n-7)	-	-	0.07	0.01
C20:0	0.22	0.03	0.18	0.02
C20:1 (n-9)	1.54	0.27	0.89	0.14
C20:3	-	-	0.07	0.01
C20:4 (n-6)	-	-	0.35	0.10
C22:4 (n-6)	-	-	0.08	0.02
C22:5 (n-3)	-	-	0.07	0.02
SFA	35.87	2.39	39.10	3.32
MUFA	55.30	2.10	56.50	3.10
PUFA	8.83	0.57	4.40	0.71

Para la toma del peso al sacrificio se utilizaron las instalaciones de Julián Martín S.L. en el matadero de Mozarbez (Salamanca) y el análisis del perfil de ácidos grasos se llevó a cabo en la Estación Tecnológica de la Carne de Guijuelo del ITACyL. Los sacrificios

3. MATERIALES Y MÉTODOS

fueron realizados en un único lote. El peso de la canal fue de 127 kg (ES = ± 12.04) El perfil de ácidos grasos fue medido bajo las mismas condiciones que en la población 1. Los ácidos grasos medidos tanto en grasa subcutánea como en grasa dorsal se muestran en la tabla 3.4.

Tabla 3.4. Perfil de ácidos grasos analizado en grasa subcutánea e intramuscular en cerdos Duroc x Ibérico del ensayo 2.

Carácter	Grasa Subcutánea		Grasa Intramuscular	
	Media	Desv. típica	Media	Desv. típica
C10:0	-	-	0.10	0.01
C12:0	0.08	0.09	0.08	0.01
C14:0	1.41	0.09	1.45	0.12
C16:0	23.60	0.81	25.51	1.58
C16:1	2.62	0.32	3.92	0.55
C17:0	0.23	0.04	0.11	0.02
C17:1	0.24	0.04	0.13	0.03
C18:0	10.66	0.95	11.68	1.23
C18:1 (n-9)	49.84	1.31	46.25	1.62
C18:1 (n-7)	-	-	4.20	0.49
C18:2 (n-6)	9.46	0.88	4.54	0.77
C18:2 (n-3)	0.70	0.07	0.25	0.04
C18:3 (n-7)	-	-	0.09	0.01
C20:0	0.20	0.03	0.19	0.02
C20:1 (n-9)	0.95	0.11	0.79	0.08
C20:3	-	-	0.08	0.02
C20:4 (n-6)	-	-	0.46	0.17
C22:4 (n-6)	-	-	0.08	0.02
C22:5 (n-3)	-	-	0.09	0.02
SFA	36.19	1.64	39.12	2.77
MUFA	53.65	1.48	55.29	2.38
PUFA	10.16	0.95	5.59	1.09

3. MATERIALES Y MÉTODOS

3.1.3. Muestras biológicas

Se tomaron muestras de sangre en tubos con EDTA (Becton Dickinson Co, Franklin Lakes, NJ) de cada uno de los animales del cruce IBMAP así como de los dos ensayos comerciales para una posterior extracción de ADN.

En el caso de los individuos pertenecientes al RC2 se obtuvieron muestras de diversos tejidos biológicos: hipotálamo, hipófisis, hígado, grasa dorsal, músculo *longissimus dorsi* y diafragma. Estos tejidos fueron conservados en nitrógeno líquido (-196°C) en el matadero y después se almacenaron en congeladores a -80°C.

3.2. EXTRACCIÓN DE ÁCIDOS NUCLEICOS

3.2.1. Extracción de ADN

La extracción se realizó a partir de 3 ml de sangre según un protocolo de extracción estándar (Sambrook *et al.*, 1989). Se añadieron 6 ml de tampón de lisis (TKM1 (10 mM TRIS HCl pH=7.6M, 10 mM KCl 1M, 10 mM MgCl₂ 1M, 2 mM EDTA 0.5 M pH=8) con Tritón 2%). Después de mezclar por inversión se centrifugó a 1200 r.p.m a 4°C durante 7 minutos. El sobrenadante fue descartado. Este paso se repitió dos veces pero sólo añadiendo el TKM1. El sedimento fue resuspendido en 6 ml de tampón de lisis de células (400 mM NaCl, 10mM TRIS-HCl pH=8.5, 5mM EDTA, 0.2% SDS) con 300 µg de proteinasa K y 12 µg de ARNasa. La mezcla se incubó a 37°C con agitación durante toda la noche para permitir la liberación del ADN del núcleo. Pasado este tiempo, el ADN fue precipitado con etanol absoluto (2.5 volúmenes) y posteriormente lavado con etanol al 80%, se mantuvo durante 15 minutos a -20°C y se centrifugó a 1200 rpm a una temperatura de 4°C durante 10 minutos. Después del centrifugado, se volvió a desechar el sobrenadante y el sedimento fue rehidratado con 300 µl de TE a pH=8 (TRIS 10 mM pH=8, EDTA 1 mM pH=8).

Las muestras de ADN extraídas fueron cargadas en un gel de agarosa al 0.8% en TBE que fue sometido a electroforesis con el fin de comprobar la integridad de las mismas. Posteriormente el ADN se cuantificó por espectrofotometría mediante un NanoDrop ND-100 (NanoDrop Technologies, Wilmington, USA). Las muestras fueron almacenadas a una temperatura de -20°C.

3. MATERIALES Y MÉTODOS

3.2.2. Extracción de ARN

Se realizaron extracciones de ARN total de hígado, músculo *Longissimus dorsi* y grasa dorsal a partir de muestras de entre 50-100 mg. Se utilizó el KitTM RiboPure (Ambion, Applied Biosystems, Foster City, USA) siguiendo las instrucciones proporcionadas por el fabricante.

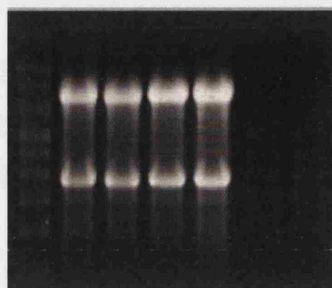
El tejido se homogenizó en 1 ml de *Tri Reagent* con el homogeneizador *workcenter* modelo *IK4-T 10 basic* (IKA-WERKE, Staufen, Alemania). La mezcla se incubó durante 5 minutos a temperatura ambiente y se realizó una centrifugación a 12.000 g durante 10 minutos a una temperatura de 4°C para eliminar el material insoluble de la mezcla. El sobrenadante fue transferido a un nuevo tubo y se añadieron 200 µl de cloroformo, se mezcló por inversión y agitación por vortex a velocidad máxima durante 15 segundos. La mezcla se incubó durante 5 minutos a temperatura ambiente y posteriormente se centrifugó en las mismas condiciones que en el paso anterior. Como resultado la mezcla se separó en tres fases: una fase inferior de color rojo (orgánica), una interfase y una fase incolora superior que es la que contenía el ARN. 400 µl de esta fase acuosa fueron transferidos a un nuevo tubo a los que se añadieron 200 µl de etanol al 100%, esta mezcla fue agitada mediante vortex a máxima velocidad durante 5 segundos para evitar la precipitación del ARN. La mezcla se transfirió a una columna del kit de extracción y fue centrifugada durante 30 segundos a temperatura ambiente a 12.000 g. Como resultado, el fluido pasó por la columna y el ARN quedó retenido en el filtro. La columna fue lavada dos veces con un tampón de lavado que contenía etanol mediante dos centrifugaciones con las mismas condiciones que en el paso anterior. Una vez realizados los lavados, el ARN fue eluido en 100 µl de tampón de elución.

3.2.2.1. Determinación de la calidad del ARN y cuantificación.

Al igual que el ADN, el ARN fue cuantificado en un NanoDrop ND-100 (NanoDrop Technologies, Wilmington, USA). Las muestras se sometieron a electroforesis en un gel de agarosa al 1.5% en TBE. Se desnaturalizó 1 µl de ARN de cada una de las muestras con 5 µl de formamida a 65°C. La calidad de ARN obtenido fue estimado mediante la observación de las intensidades relativas de las bandas correspondientes a los ARN ribosomales 28S y 18S (Figura 3.2). Se considera una buena calidad cuando la relación entre la intensidad de estas bandas es aproximadamente igual a 2.

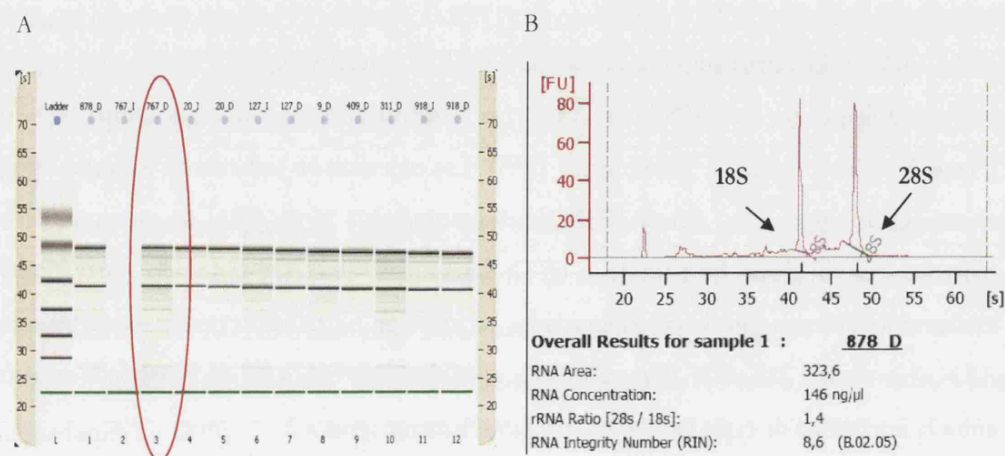
3. MATERIALES Y MÉTODOS

Figura 3.2. Gel de ARN



Por otro lado, para realizar estudios de expresión mediante PCR cuantitativa y microarrays es necesaria una medida más precisa de la calidad de las muestras para evitar la subestimación del ARN expresado debido a su degradación. Para ello las muestras fueron analizadas en un Bioanalizador ARN Nano 6000 Labchip kit mediante un Bioanalizador Agilent 2001 (Agilent Technologies, Palo Alto, USA) (Figura 3.3). Para ello, las muestras se cargan en un LabChip y se realiza una electroforesis a través de una red de canales en miniatura sobre un polímero de cristal, plástico o silíceo. La muestra va pasando por estos canales y se va separando en diferentes componentes: ARN 28S, ARN 18S, ARN5S y ARN5'8S, ARN de transferencia y ARNs mensajeros. El resultado es procesado con el programa 2100 Expert Software v.B.02.05 generando una imagen que simula un gel de electroforesis (Figura 3.3). En el presente trabajo se han utilizado los chips *Eukariote total RNA Nano Series II versión 2.5*.

Figura 3.3. A) Simulación gráfica de un gel de electroforesis de agarosa realizada con el programa 2100 Expert Software. B) Perfil de absorbancia de la muestra número 1 de la figura 3.3.A. con los datos de calidad (RIN) y concentración.



3. MATERIALES Y MÉTODOS

El programa analiza las áreas de los picos correspondientes a los ARNr 28S y 18S y calcula el ratio 28S/18S. Considerando las áreas de los fragmentos de diferente tamaño el programa calcula el *ARN Integrity Number* (RIN) que cuantifica la integridad del ARN y cuyo valor máximo es igual a 10. Todas las muestras utilizadas en este trabajo para los análisis de expresión génica mediante PCR cuantitativa mostraron un RIN entre 7 y 10.

3.3 ANÁLISIS MOLECULAR DE GENES CANDIDATOS

3.3.1. Caracterización, secuenciación y búsqueda de polimorfismos.

Se analizó la secuencia codificante de tres genes candidatos: *Acyl-Coenzyme A oxidase 1* (*ACOX1*), *ATP citrate lyase* (*ACLY*) y *Sterol regulatory element binding factor 1* (*SREBF1*) en dos animales Guadrybas, dos Landraces; en el caso del gen *SREBF1* se analizaron también dos individuos RC2.

A partir de 2 µg de ARN total se sintetizó el ADN copia en un volumen total de 20 µl que contenía 200 ng de hexámeros aleatorios, 500 µM de dNTPs, tampón 5X "First Strand" (Invitrogen), 10 mM de DTT, 40 U de RNAsin (Promega) que actúa como inhibidor de ARNasas y 200 U de transcriptasa inversa SuperScript™ II (Invitrogen, Life Technologies). Las incubaciones se realizaron siguiendo el protocolo indicado por el fabricante (Invitrogen, Life Technologies) en un termociclador PTC-100 (MJ, Research, Waterton, MA).

Para poder amplificar la región codificante de los genes candidatos se realizó la búsqueda de las secuencias codificantes en las bases de datos disponibles (<http://www.ncbi.nlm.nih.gov/> y <http://www.ensembl.org/>). A partir de las secuencias encontradas se diseñaron las parejas de oligonucleótidos para hacer las amplificaciones utilizando el programa QUICKPRI del paquete informático Winstar. En el diseño de los oligonucleótidos se tuvieron en cuenta los siguientes parámetros: un tamaño de 17-24 bases, valores de Tm entre 48°C y 70°C y evitar la formación de dímeros y horquillas superiores a 3 pb.

Las reacciones de amplificación se llevaron a un volumen final de 25 µl conteniendo 2 µl de ADNc, tampón estándar 1X (75mM Tris-HCl pH=), 50 mM KCl, 20 mM (NH₄)₂SO₄, 0.2 mM de dNTPs, 0.5 µM de oligonucleótidos y 0.5 U de polimerasa Tth (Biotools). La concentración de MgCl₂ fue optimizada para cada fragmento. Para la puesta a punto de la temperatura óptima de hibridación de cada reacción, fue utilizado un

3. MATERIALES Y MÉTODOS

termociclador de gradiente PTC-200 (MJ, Research , Waterton, MA). El perfil térmico empleado consistía en un periodo de desnaturalización a 94°C durante 5 minutos, seguido de 40 ciclos a 94°C 30 segundos, 45 segundos a la temperatura de hibridación específica de cada pareja de oligonucleótidos y 45 segundos a 72°C con un periodo de elongación final de 10 minutos a 72°C. El resto de las amplificaciones se llevaron a cabo en un termociclador automático PTC-100 (MJ Research, Waterton, MA).

Los productos de amplificación fueron verificados mediante electroforesis en gel de agarosa al 2.5% en TBE y posteriormente teñidos en bromuro de etidio. Los geles se observaron en un transiluminador de luz ultravioleta (UV) acoplado a un procesador de imágenes (Gel Doc 2000, BIO-RAD) para ser visualizadas mediante el programa “Quantity One 4.3.0”. El tamaño de los fragmentos amplificados se estimó mediante comparación con un patrón de tamaño molecular (HyperLadder II 50-2000 pb, ECOGEN).

Los fragmentos amplificados fueron purificados. Esta purificación se realizó directamente de la reacción de PCR cuando se obtuvo solo una banda. En el caso de observar en el gel de agarosa varias bandas inespecíficas se realizó la purificación cortando la banda de agarosa correspondiente al fragmento deseado y purificando la misma. En ambos casos se usó el kit de purificación de Amershan Biosciences (*GFX PCR DNA and Gel Band Purification kit*) siguiendo las instrucciones del fabricante y se eluyó el producto en 30 µl de agua MiliQ. 1 µl de cada producto purificado fue cargado en geles de agarosa al 2.5% en TBE con el objeto de comprobar la purificación y para realizar una cuantificación del mismo mediante comparación con patrones de tamaño molecular de concentración conocida. Esta cuantificación es necesaria ya que para realizar la secuenciación de productos de menos de 500 pb se requería una concentración de 10 µg/µl y para productos con tamaño comprendido entre 500 y 100 pb se requerían concentraciones de 20 µg/µl.

Los productos purificados fueron secuenciados en ambas direcciones con el 3100 BigDye® Terminator v3.1 Matrix Standard kit en un secuenciador multicapilar 3730 DNA Analyzer (Applied Biosystems) de los servicios de secuenciación de la empresa Secugen (empresa *spin-off* del Centro de Investigaciones Biológicas, Madrid).

Las secuencias de los fragmentos amplificados fueron editadas mediante el programa EDITSEQ (WinStar). Una vez editados, los fragmentos fueron ensamblados para poder obtener las secuencias completas de cada uno de los ADNc. Por último, con el fin de detectar polimorfismos, las secuencias ensambladas de cada uno de los individuos se alinearon mediante el programa MEGALIGN (WinStar).

3. MATERIALES Y MÉTODOS

Una vez detectados los polimorfismos se analizaron las posibles implicaciones biológicas que pueden tener estos cambios. En primer lugar se verificó si estos cambios nucleotídicos producían cambio aminoacídico o por el contrario eran silentes. Si el cambio era en la secuencia de la proteína se analizó el posible efecto que podría tener en la función de la misma mediante distintas herramientas bioinformáticas:

- SIFT (*Sorting Intolerant From Tolerant*, sift.jcvi.org/) genera una predicción sobre el efecto que un cambio aminoacídico podría tener en la función de la proteína mediante comparación de las diferentes secuencias de una proteína disponibles en las bases de datos. En función del efecto sobre la función, estos cambios son clasificados como tolerados o no tolerados (Ng y Henikoff, 2006).
- SMART (*Simple Modular Architecture Research Tool*, smart.embl-heidelberg.de/) analiza si los cambios aminoacídicos detectados se encuentran en un dominio funcional de la proteína y si tienen algún efecto (Letunic *et al.*, 2006).
- PSIPRED (*Protein Structure Prediction Server*, bioinf.cs.ucl.ac.uk/psipred/) infiere si el cambio aminoacídico produce algún cambio de conformación en la estructura secundaria de la proteína.
- RNA *Secondary Structure Prediction* (genebee.msu.su/services/rna2_reduced.html) infiere si los SNPs dan lugar a posibles cambios estructurales en la conformación secundaria del RNA.

También se analizó el posible efecto que podrían tener las mutaciones silentes en la función biológica mediante:

- TargetScan versión 5.1 (targetscan.org) predice si, en una secuencia problema, existen dianas de unión de microARNs (miARNs). Para ello hace una búsqueda de octámeros, heptámeros y hexámeros conservados y no conservados que sean homólogos a la región de unión de cada microARN (Friedman *et al.*, 2009).
- Pictar (pictar.mdc-berlin.de/) está basado en un algoritmo que posibilita la predicción de secuencias diana de miARNs mediante el alineamiento de secuencias 3'UTR ortólogas y busca un set de miARNs coexpresados. Después mapea los sitios diana y los filtra según la energía libre predicha (Krek *et al.*, 2005).

A pesar de que las herramientas asociadas a la búsqueda de secuencias dianas de microARNs en los extremos 3'UTR de los genes están basadas principalmente en bases de datos de secuencias de humano, ratón o vacuno entre otros, éstas pueden tenerse en cuenta debido a la elevada homología de estas secuencias dianas y de los microARNs

3. MATERIALES Y MÉTODOS

entre las especies anteriormente mencionadas. Estas dos herramientas tienen una función similar, pero es más eficaz utilizar ambas y comparar los resultados obtenidos.

3.3.2. Genotipado de polimorfismos

La selección de polimorfismos a genotipar se realizó combinando criterios de relevancia biológica y viabilidad técnica. Se dio preferencia a aquellos que originan cambio aminoacídico frente a los silentes, así como a los localizados en posibles regiones reguladoras. La técnica básica de genotipado fue la pirosecuenciaciónTM (Ronaghi, 2003), aunque la técnica de PCR-RFLP se utilizó para genotipar el polimorfismo *ACACA*: g.48001 A>G

3.3.2.1. PCR-RFLP

Se diseñó una pareja de oligonucleótidos para amplificar mediante PCR un fragmento de 500 pb que contiene el polimorfismo *ACACA*: g.48001 A>G. Este polimorfismo genera dianas polimórficas de restricción reconocidas por la enzima de restricción *MspI*. Las reacciones de amplificación fueron llevadas a un volumen final de 25 µl (condiciones en apartado 3.3.1). Una vez realizada la amplificación se llevaron a cabo las reacciones de digestión en un volumen final de 20 µl, conteniendo 5 µl de producto de la PCR, 10 U de la enzima *MspI* (New England Biolabs) y 2.5 µl de buffer de digestión 10X específico de la enzima (NEB 2). Estas reacciones fueron incubadas a 37°C durante toda la noche.

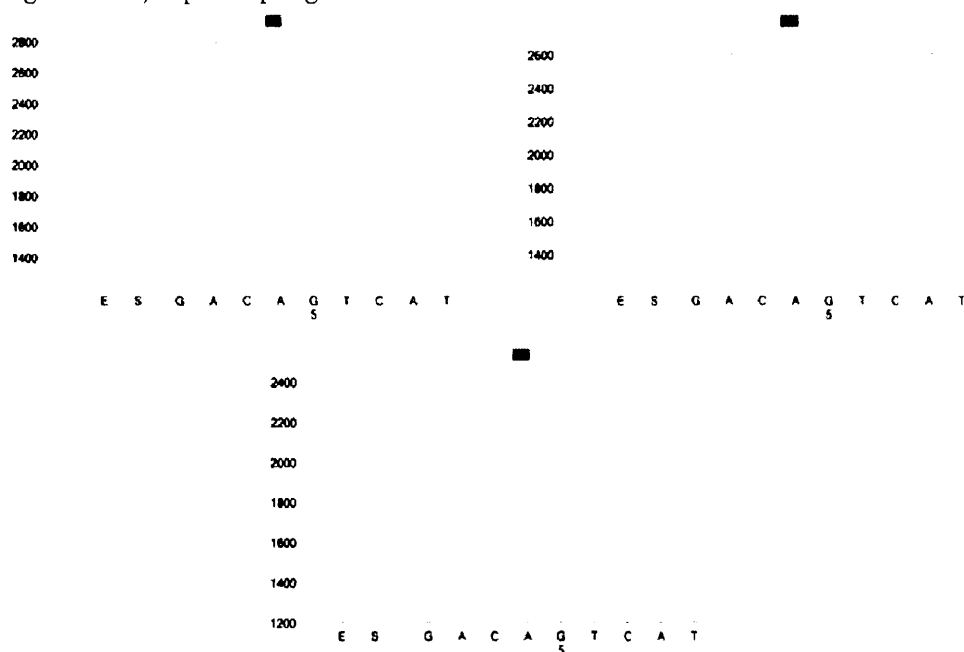
3.3.2.2. PirosecuenciaciónTM

Esta técnica está basada en un principio de secuenciación por síntesis mediante el uso de un complejo enzimático, se trata de una minisequenciación a tiempo real. En un primer paso se amplifica por PCR la región de interés que contiene el polimorfismo a genotipar. Para ello se diseña una pareja de oligonucleótidos con el programa PSQ Assay Design (Biotage AB) versión 1.0.6, teniendo en cuenta que la longitud de los fragmentos amplificados no debe superar los 200 pb. Uno de los dos oligonucleótidos empleados en la PCR debe estar marcado con biotina en el extremo 5' para posibilitar el posterior aislamiento de fragmentos de cadena sencilla. Se desnaturalizan los fragmentos de PCR y se capturan las hebras marcadas con biotina mediante la utilización de esferas de sefarosa recubiertas por estreptavidina. En un segundo paso se produce la reacción de

3. MATERIALES Y MÉTODOS

pirosecuenciación propiamente dicha, en donde se emplea un tercer oligonucleótido diseñado conjuntamente con la pareja de cebadores utilizada en la reacción de amplificación. El extremo 3' de este oligonucleótido hibrida en una posición adyacente al polimorfismo, a partir de la cual se produce la extensión de la secuencia mediante la acción de la polimerasa contenida en el complejo enzimático que cataliza la adición de nucleótidos de manera secuencial. En el mismo complejo están presentes también las enzimas luciferasa, sulfurilasa y apirasa. Las dos primeras permiten convertir los productos generados durante la incorporación de nucleótidos (pirofosfato) en luz que es detectable por un dispositivo CCD y la apirasa degrada el ATP y los nucleótidos no incorporados. Los nucleótidos se dispensan por separado y en un orden conocido y la detección de la luz emitida permite inferir el número de nucleótidos incorporados en la reacción de secuenciación.

Figura 3.4. Ejemplo de pirograma



La reacción de amplificación se realizó en un volumen final de 25 μ l empleando las condiciones descritas en el apartado 3.3.1. Las reacciones de PCR se diluyeron con agua MilliQ autoclavada en un volumen final de 40 μ l, la cantidad de reacción de PCR diluida varió en función de la intensidad de la banda visualizada en el gel de agarosa. A esta dilución se añadió 2 μ l de Streptavidin SepharosaTM HP (Amersham Biosciencie) y 38 μ l de tampón de unión, pH=7. La mezcla se incubó a temperatura ambiente con agitación durante 5 minutos para facilitar la unión de la sepharosa con los fragmentos biotinilados. Paralelamente se añadió en una placa de pirosecuenciación (Biotage) 12 μ l de dilución de

3. MATERIALES Y MÉTODOS

oligonucleótido de secuenciación (3M) en tampón de hibridación (pH=7.6). Después se colocó la placa con la PCR en la estación de trabajo de vacío y se filtró la mezcla con el objetivo de retener y limpiar los fragmentos de cadena sencilla biotinilados. Posteriormente se mezcló el producto de PCR purificado con el oligonucleótido diluido y se incubó a 80°C durante 2 minutos para favorecer la hibridación.

La mezcla se introdujo en el pirosecuenciador donde se realizó la secuenciación del fragmento. Como resultado el programa calcula las correlaciones entre los picos de luz y los nucleótidos añadidos en cada momento genera los correspondientes pirogramas donde se representa la emisión de luz emitida al añadir cada nucleótido y se puede distinguir el genotipo de cada muestra (Figura 3.4).

3.3.3. Detección de posibles secuencias promotoras con herramientas bioinformáticas.

Se utilizaron distintas aplicaciones disponibles en la red con el objetivo de detectar secuencias con motivos de unión a factores de transcripción.

- TFSEARCH.(*Transcription Factor Search*) - Se introduce la presunta secuencia promotora en la aplicación y ésta realiza una búsqueda de la misma en la base de datos TRANSFAC. Esta base de datos contienen secuencias reguladoras de ADN reguladora y factores de unión a las mismas. Al introducir la presunta secuencia regulatoria la aplicación da información diferente acerca de: los diferentes motivos de unión a factores de transcripción detectados en la secuencia introducida, una pequeña explicación de la función del gen introducido, de la distribución celular de los factores de transcripción para los que se ha detectado motivos de unión y las clases de los factores de transcripción (www.cbrc.jp/research/db/TFSEARCH.html) (Heinemeyer *et al.*, 1998).
- Genomatix Software Suite. (Cartharius *et al.*, 2005) Se usó el módulo MatInspector Release (Genomatix Software GmbH, München, Alemania) Esta aplicación utiliza una librería de matrices descriptivas para lugares de unión a factores de transcripción para localizar posibles factores de transcripción dentro de la secuencia de ADN introducida en la aplicación. Ésta además ofrece información acerca de los factores de transcripción, sitios de unión y dominios.

3. MATERIALES Y MÉTODOS

3.4. PCR CUANTITATIVA A TIEMPO REAL

3.4.1. Diseño de oligonucleótidos

La técnica de PCR cuantitativa a tiempo real se empleó para cuantificar la expresión del gen *ACACA*. Los oligonucleótidos empleados en la técnica de cuantificación por qPCR mediante SYBR Green fueron diseñados mediante el programa QUICKPRI del paquete Winstar. Previamente al diseño de los mismos se realizó un estudio pormenorizado del gen (número de exones y transcritos descritos) para determinar que regiones eran más adecuadas para amplificar en función del transcrito a detectar. También se tuvieron en cuenta otros aspectos como evitar regiones que pudiesen formar estructuras secundarias o largas repeticiones del mismo oligonucleótido. Además se realizó el diseño en regiones exclusivas del fragmento de interés y abarcando más de un exón para evitar la amplificación de ADN genómico residual. Por último el tamaño de los fragmentos amplificados variaba entre 100 a 400 pb.

Las secuencias utilizadas para el diseño de oligonucleótidos fueron obtenidas de la base de datos del NCBI (www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Nucleotide) y la información relativa a la estructura de los transcritos de la base de datos Ensembl (www.ensembl.org/index.html).

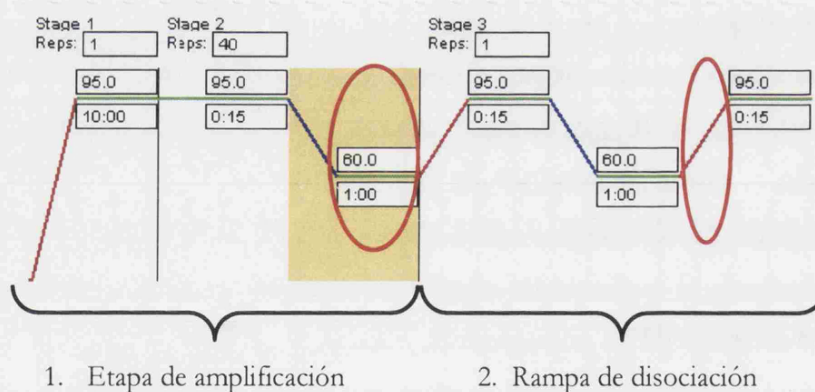
3.4.2. Protocolo

En una primera etapa se realizaron PCRs convencionales con ADNc para verificar si las parejas de oligonucleótidos amplificaban de manera específica y el tamaño de los fragmentos correspondía al esperado. El ADNc fue sintetizado tal y como se describe en el apartado 3.3.1. Una vez comprobada la funcionalidad de los oligonucleótidos se realizó la cuantificación de los transcritos en un termociclador MX3000P (Stratagene) usando SYBR Green mix de Takara. Las reacciones de PCR se llevaron a un volumen final de 20 µl que contenía 2.5 µl de ADNc (dilución 1/20), 10 µl de SYBR Green, 0.15 mM de cada oligonucleótido y 0.4 µl del fluoróforo de referencia ROX que se utilizó como referencia interna para normalizar la emisión de fluorescencia propia de los pocillos. El perfil térmico de la reacción de amplificación consistía en un periodo de activación de la polimerasa de 10 minutos a 95°, seguido por 40 ciclos de 95° durante 15 segundos y 60° durante 1 minuto. La fluorescencia fue registrada al final de la etapa de 60° en cada uno de los 40 ciclos.

3. MATERIALES Y MÉTODOS

Se incluyeron controles negativos y positivos para todos los genes testados. Además se realizó una Curva de Disociación del producto amplificado en cada reacción. Para ello se aplicó un gradiente creciente de temperaturas después de la reacción de la PCR y se monitorizó la cinética de disociación de los fragmentos amplificados. La curva se obtuvo mediante la adición de un ciclo de 95° durante 15 segundos, 60° durante 1 minuto y una rampa de temperatura con incrementos de 0.01°C/segundo hasta alcanzar los 95°C (Figura 3.5). La curva de disociación indica la temperatura de fusión de cada fragmento amplificado en la reacción. Al ser específica para cada uno de ellos permite verificar si existe más de un fragmento amplificado en la reacción, analizar mutaciones puntuales e incluso observar la formación de dímeros de oligonucleótido.

Figura 3.5. Condiciones de amplificación de la PCR cuantitativa en tiempo real. Los pasos señalados indican los momentos de la reacción en los que se mide la fluorescencia

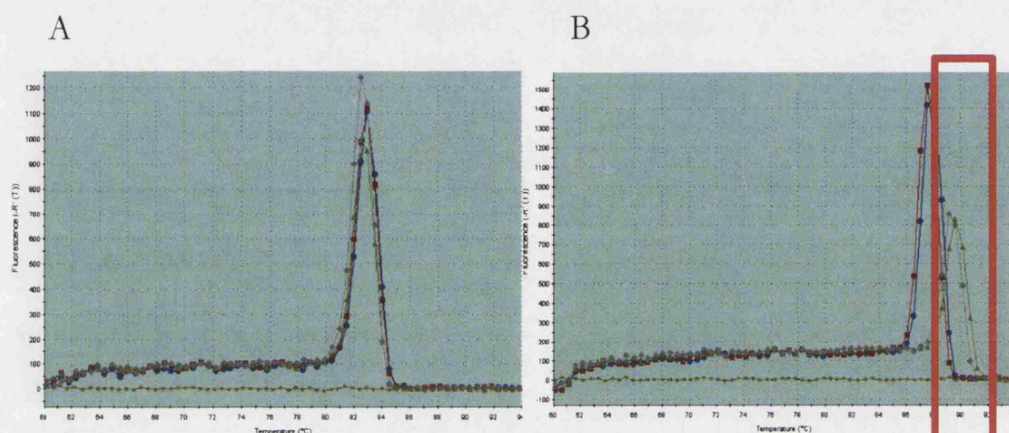


Antes de realizar las cuantificaciones de los transcritos hay que optimizar la concentración de los oligonucleótidos de manera que produzcan el máximo número de transcritos y sin que no se formen dímeros de oligonucleótido. Para ello se cuantificó un mismo ADNc utilizando diferentes concentraciones de oligos.

Las cuantificaciones se realizaron por triplicado para reducir al máximo el error técnico del láser. Los datos obtenidos se analizaron mediante el software MxPro (Stratagene) que permite calcular de manera automática la media de las réplicas y el Ciclo umbral (Ct) de las muestras teniendo en cuenta el fluoróforo de referencia.

3. MATERIALES Y MÉTODOS

Figura 3.6. A). Curva de disociación correcta. B). Curva de disociación con amplificación inespecífica

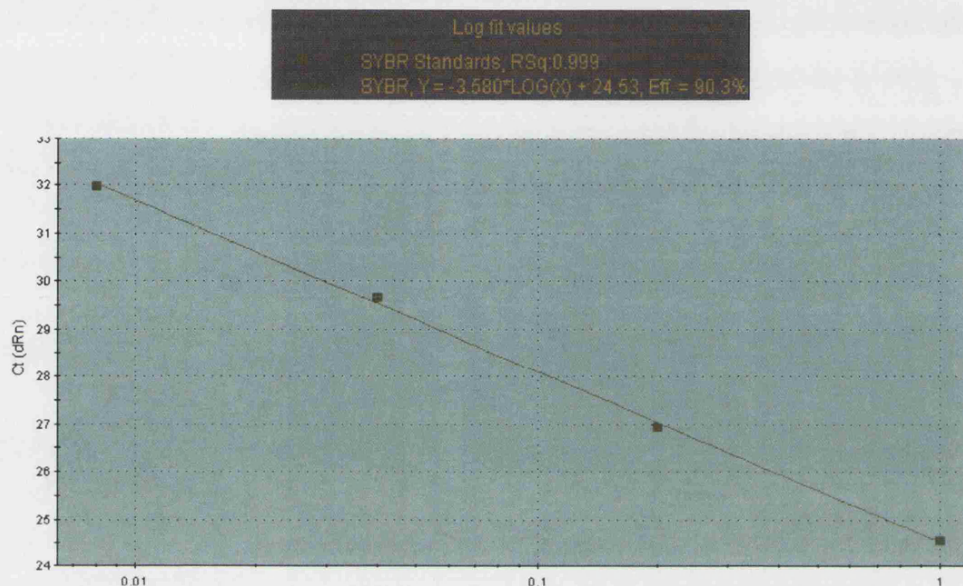


3.4.3 Curva estándar

Antes de cuantificar la expresión de un gen hay que valorar la eficiencia, sensibilidad y precisión de la PCR en el tejido a estudiar. Para realizar esta valoración se hizo una dilución seriada de ADNc con 4 puntos (1, 1/5, 1/25 y 1/125) y se cuantificó la expresión de los transcritos. Posteriormente se representó en un gráfico el logaritmo de la cantidad de molde frente a los valores Ct de las diluciones seriadas obteniéndose una Curva Estándar (Figura 3.7). Para incrementar la precisión las diluciones fueron analizadas por triplicado. La pendiente (m) de la Curva Estándar es utilizada para calcular la eficiencia de la reacción. Según el manual de Stratagene los valores de m óptimos deben estar comprendidos entre -3.1 y -3.6. Si la eficiencia fuese del 100% en cada uno de los ciclos la cantidad de molde debería duplicarse en cada ciclo debido a que la PCR es una reacción exponencial. Sin embargo la mayoría de las reacciones PCR no alcanzan la eficiencia óptima y hay que usar un factor de corrección apropiado que no sobreestime la cantidad de ADNc de la que partimos. Este factor denominado eficiencia de la reacción (E) se calcula como $E = [10^{(-1/m)}] - 1$ (Wong y Medrano, 2005). Fueron aceptadas como válidas las eficiencias superiores al 85%.

3. MATERIALES Y MÉTODOS

Figura 3.7. Ejemplo de curva estándar de la reacción de PCR cuantitativa



3.4.4. Obtención de resultados

Para analizar los datos obtenidos se realizó un método de cuantificación relativa utilizando el *comparative Ct* (Livak y Schmittgen, 2001). Este método se basa en un modelo matemático que calcula los cambios de expresión génica mediante la referencia relativa entre la muestra y el gen control. Este modelo contempla la inclusión de un factor de corrección para la eficiencia de la amplificación. De tal manera que los Ct fueron transformados en cantidades relativizando la expresión de cada muestra de un gen al de mayor expresión (menor Ct) mediante la siguiente expresión:

$$Q \text{ (cantidad de molde)} = 10^{-\Delta Ct / \text{pendiente}}$$

Para aplicar la relativización se usó el programa geNorm (medgen.ugent.be/~jvdesomp/genorm/) (Vandesompele *et al.*, 2002). Esta herramienta toma como factor de normalización la media geométrica entre pares de genes controls que muestran un perfil de expresión similar, partiendo del principio de que pares de genes que presenten patrones de expresión estables entre sí, son genes control adecuados.

3.4.5. Normalización de resultados

En este trabajo se han utilizado dos genes control para la normalización de resultados. Los genes seleccionados fueron, *GAPDH* y el gen *B2M*, esta selección se basó en una búsqueda bibliográfica. Las secuencias de las parejas de oligonucleótidos fueron obtenidas de Kuijk *et al.* (2007). En cada placa de PCR, y para cada muestra problema, fueron amplificados por triplicado, tanto el gen problema como dos genes control. Se seleccionó el ciclo umbral en el punto en el que la pendiente de la PCR (m) es máxima y se tuvo en cuenta el mismo valor para todos los ensayos de un mismo gen problema. Además, se incluyó un control positivo para poder comparar resultados de diferentes placas.

Una vez se obtuvo el Ct de cada muestra, se calculó la cantidad relativa de molde (Q) de la muestra problema y de los genes control. Los datos de los control fueron utilizados para calcular un factor de normalización mediante el programa geNorm. Después de aplicar este factor se obtuvieron los datos de expresión normalizados.

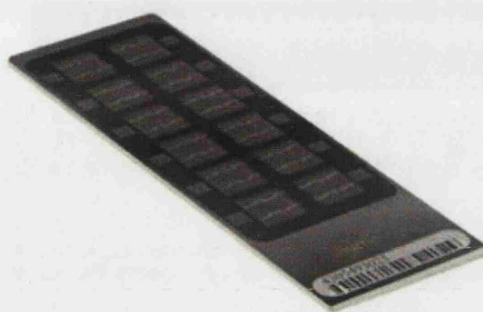
3.4.6. Análisis estadístico de resultados

Una vez normalizados los datos, se realizó un análisis de mínimos cuadrados en el cual se testa la expresión de cada gen teniendo en cuenta tres niveles de cada genotipo (homocigoto 1, homocigoto 2 y heterocigoto). Mediante esta técnica se analizó la expresión del gen *ACACA* en grasa dorsal en función del genotipo de los polimorfismos *ACACA* c.5634T>C y *ACACA* g. 48001 A>G, los datos de expresión fueron corregidos por el lote de sacrificio y por el sexo. Estos análisis fueron realizados con el paquete estadístico SAS.

3.5 GENOTIPADO A GRAN ESCALA

Para realizar el genotipado a gran escala se utilizó la plataforma comercial *PorcineSNP60 Genotyping BeadChip* (Ramos *et al.*, 2009) de Illumina (San Diego, USA). Este chip consiste en una plataforma sólida en la que se encuentran inmovilizadas 64.232 sondas de una longitud aproximada de 50 bases que hibridarán con fragmentos del ADN genómico a genotipar con los que tengan una alta homología (Figura 3.8). Mediante esta técnica se genotiparon 427 individuos que comprendían los individuos de las generaciones F3, RC1 y RC2 y sus respectivos ascendientes procedentes de las generaciones F0, F1 y F2.

Figura 3.8. Fotografía del *PorcineSNP60 Genotyping BeadChip*



3.5.1. Cuantificación, hibridación y escaneado

El proceso de cuantificación, hibridación y escaneado se llevó a cabo en el *Servicio Veterinario de Genética Molecular* (Universitat Autònoma de Barcelona, Spain). En primer lugar, la cuantificación del ADN de las muestras problema fue realizada mediante el protocolo Quant-iT™ (Invitrogen, Life Technologies), este método de cuantificación es preferido frente a los métodos tradicionales de absorbancia ultravioleta ya que es más preciso. La medida está basada en la emisión de fluorescencia de un fluoróforo que se une de manera específica a la molécula de ADN sin sufrir interferencia de nucleótidos libres u otros contaminantes. La hibridación se realizó siguiendo el protocolo estándar *Infinium r II Assay* de Illumina (San Diego, USA) basado en la amplificación del DNA y posterior fragmentación por un proceso enzimático controlado que no requiere electroforesis en gel. A continuación tiene lugar la hibridación de las sondas correspondientes a cada uno de los SNPs en posición adyacente al polimorfismo seguida de la unión específica de nucleótidos marcados con fluorescencia. Para llevar a cabo la hibridación se utilizaron 750 ng de DNA genómico total. (Figura 3.9). Una vez hibridados, los chips fueron escaneados en el equipo *Bead Station 500* (Figura 3.10).

3. MATERIALES Y MÉTODOS

Figura 3.9. Protocolo *Infinium II Assay* de Illumina para el procedimiento de hibridación y escaneado

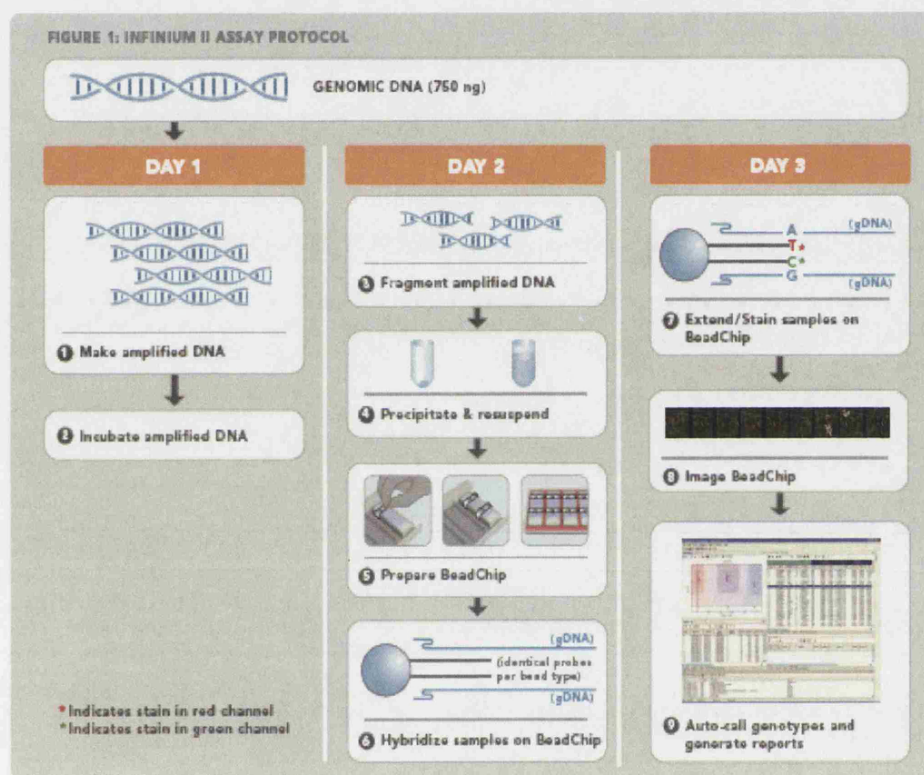


Figura 3.10. Fotografía del equipo de escaneado *Bead Station 500*



3. MATERIALES Y MÉTODOS

3.5.2. Extracción y procesamiento de datos de genotipado

Una vez escaneados los chips, el equipo generó archivos de intensidad de datos (*.idat), de los cuales se extraen los datos de genotipado. Para llevar a cabo este proceso se utilizó el programa *GenomeStudio Genotyping Module* (Illumina, San Diego, USA). Se filtraron aquellos SNPs que no superaban un Umbral de Llamadas (Gene Call Threshold) de 0.15 y se tuvo que tener en cuenta el Índice de Llamadas (Call Rate) de cada muestra de manera individual. Este índice depende del número de SNPs que se han podido clasificar en homocigotos y heterocigotos. Todas las muestras presentaron un Índice superior a 0.99.

Después de verificar la calidad del genotipado los datos fueron exportados en archivos *.csv. Posteriormente las sondas fueron anotadas según los datos publicados en la base de datos Ensembl (build9) (www.ensembl.org/Sus_scrofa/Info/Index). Del total de las sondas genotipadas se seleccionaron aquellas pertenecientes al SSC12, mediante el programa PLINK (Purcell *et al.*, 2007). Se seleccionaron aquellos SNPs que presentaron un porcentaje de genotipos faltantes inferior a 10% y en los que el alelo de menor frecuencia (Minor Allele Frequency, MAF) tenía una frecuencia superior a 0.05. Después de este filtrado se obtuvieron 1197 SNPs genotipados.

3.6. ANÁLISIS ESTADÍSTICOS

3.6.1. Mapeo de ligamiento

Los mapas de ligamiento se construyeron usando la opción BUILD del programa CRI-MAP versión 2.4 (Green *et al.*, 1990). Esta opción está basada en la función de mapeo de Kosambi la cual se fundamenta en la proporción de dobles sobrecruzamientos respecto de la distancia y ajusta estas distancias teniendo en cuenta el fenómeno de interferencia que postula que la probabilidad de que se produzca un sobrecruzamiento en un punto disminuye si ya se ha producido un sobrecruzamiento en el mismo segmento. Por lo tanto, el programa calcula la distancia entre loci (m) en función de la frecuencia de recombinación (r) mediante la ecuación:

$$m = 1/4 \ln(1+2r) - \ln(1-2r)$$

Se construyeron tres mapas de ligamiento diferentes:

3. MATERIALES Y MÉTODOS

- Mapa 1: Construido a partir de datos de genotipado de 6 marcadores tipo microsatélites (*s0106*, *s0143*, *sw874*, *sw1307*, *swr1021* y *swr1802*) y de 7 SNPs detectados en sus respectivos genes candidatos (*ACACA*, *ACLY*, *ACOX*, *FASN*, *GH*, *GIP* y *SREBF*). Estos marcadores fueron genotipados en todas las generaciones del cruce IBMAP.
- Mapa 2: Construido a partir de los registros de los 6 microsatélites utilizados en la construcción del mapa 1, del SNP del gen candidato *SREBF1* y los registros de 26 SNPs (ver Anejo) genotipados mediante el chip de SNPs *PorcineSNP60 Beadchip* (*SNP119*, *SNP196*, *SNP273*, *SNP350*, *SNP362*, *SNP370*, *SNP371*, *SNP377*, *SNP382*, *SNP437*, *SNP599*, *SNP613*, *SNP618*, *SNP624*, *SNP635*, *SNP645*, *SNP673*, *SNP707*, *SNP796*, *SNP849*, *SNP904*, *SNP960*, *SNP965*, *SNP972*, *SNP974*, y *SNP975*). Se seleccionaron estos 26 marcadores en función de que ellos estuviesen distribuidos de manera uniforme a lo largo del SSC12 porcino y que presentasen una $MAF > 0.30$. Estos marcadores fueron genotipados en las generaciones F3, RC1 y RC2.
- Mapa 3. Construido con 50 de los 489 marcadores del chip de SNPs *PorcineSNP60 Beadchip* seleccionados para realizar la detección de QTLs debido a que la versión 2.4 del programa CRI-MAP no permitía utilizar un número tan elevado de marcadores. (*SNP1*, *SNP41*, *SNP51*, *SNP61*, *SNP71*, *SNP91*, *SNP101*, *SNP141*, *SNP151*, *SNP161*, *SNP171*, *SNP191*, *SNP211*, *SNP231*, *SNP251*, *SNP301*, *SNP231*, *SNP251*, *SNP301*, *SNP321*, *SNP371*, *SNP381*, *SNP391*, *SNP401*, *SNP431*, *SNP441*, *SNP451*, *SNP471*, *SNP561*, *SNP601*, *SNP611*, *SNP621*, *SNP631*, *SNP651*, *SNP661*, *SNP681*, *SNP691*, *SNP701*, *SNP711*, *SNP741*, *SNP761*, *SNP781*, *SNP831*, *SNP841*, *SNP851*, *SNP861*, *SNP901*, *SNP911*, *SNP921*, *SNP931*, *SNP941* y *SNP977*, ver Anejo).

3.6.2. Análisis de detección de QTLs

Los modelos de detección de QTL en cruces de líneas divergentes asumen alelos con efecto sobre el carácter (Q y q) fijados alternativamente en cada una de las líneas parentales. La primera etapa de estos análisis se ha basado en la comparación de diferentes posiciones del cromosoma (por ejemplo, cada cM) de la verosimilitud de un modelo animal que incorpora los efectos del QTL (por ejemplo, modelos $M1a$ = efecto aditivo o $M1ad$ =

3. MATERIALES Y MÉTODOS

efectos aditivo y dominante) con la de un modelo animal que corresponde a la hipótesis nula de ausencia de QTL (modelo $M0$). Estos modelos pueden representarse como sigue:

$$M0: y = X\beta + Zu + e$$

$$M1a: y = X\beta + Zu + w_1a + e$$

$$M1ad: y = X\beta + Zu + w_1a + w_2d + e$$

donde y corresponde al vector $n \times 1$ de observaciones de cada carácter analizado; β es el vector de efectos sistemáticos que incluye los del sexo y lote de sacrificio además de los de covariables como el peso de la canal; $u \sim N(0, A\sigma_u^2)$ es un vector $n \times 1$ que incluye los efectos aditivos poligénicos infinitesimales y A la matriz $n \times n$ de relaciones aditivas; w_1 y w_2 son vectores $n \times 1$ de los coeficientes c_a y c_d para los efectos fijos aditivos (a) y dominantes (d) del QTL; X y Z son las matrices de incidencia que relacionan los parámetros β y u con las observaciones, y por último e es el vector $n \times 1$ de los efectos residuales. Los coeficientes c_a y c_d se calcularon como $c_a = \text{pr}(QQ) - \text{pr}(qq)$ y $c_d = \text{pr}(Qq)$, donde $\text{pr}(QQ)$ es la probabilidad de ser homocigoto de origen Ibérico, $\text{pr}(qq)$ es la probabilidad de ser homocigoto de origen Landrace y $\text{pr}(Qq)$, la probabilidad de ser heterocigoto. Los valores de probabilidad y de los coeficientes fueron calculados para cada una de las posiciones cromosómicas analizadas, a partir de la información de los marcadores genéticos disponibles.

La detección de QTLs mediante un barrido cromosómico consiste en el cálculo de la razón de verosimilitud (LR) mediante la comparación del modelo completo con el modelo sencillo para cada posición cromosómica considerada. Se estiman los componentes de varianza del carácter fenotípico analizado (σ_u^2 , σ_e^2) así como los efectos fijos (incluyendo los del QTL) en la posición de valor máximo de LR. Los valores de la probabilidad nominal (P) de los efectos del QTL se calcularon asumiendo una distribución chi-cuadrado (χ^2_m) con un número de grados de libertad (m) equivalente a la diferencia entre el número de parámetros estimados en los modelos completo ($M1$) y reducido ($M0$), con dos grados de libertad cuando se testa aditividad y dominancia ($M1ad$) y uno cuando se testa sólo aditividad ($M1a$). La posición y magnitud de los efectos del QTL corresponden a los valores de máxima verosimilitud. Los umbrales de significación de los QTLs fueron calculados como se describe en Nezer *et al.* (2002). Los intervalos de confianza al 95% fueron calculados de acuerdo con Mangin *et al.* (1994).

3. MATERIALES Y MÉTODOS

En el caso de los análisis realizados con 489 marcadores no es correcto calcular los umbrales de significación según Nezer *et al.* (2002), debido a que la posición de los mismos está inferida a partir de la regresión lineal entre la posición física y genética de los 50 marcadores utilizados para realizar el mapa de ligamiento 3 (epígrafe 3.6.1). Por lo tanto los umbrales fueron establecidos siguiendo el criterio de corrección mediante la tasa de falso descubrimiento (FDR) propuesta por Benjamini y Yekutieli (2005) y Lander y Kruglyak (1995) estableciendo como QTL con efectos significativos sobre el carácter aquellos con un $FDR < 5\%$ ($0.05/n^\circ$ test realizados) y con efectos sugestivos, aquellos con un $FDR < 10\%$ ($0.10/n^\circ$ test realizados).

Se realizaron diferentes análisis complementarios de detección de QTL para lo que se utilizaron modelos con una estructura similar a los ya descritos: modelos con dos QTLs en diferentes posiciones del cromosoma para un mismo carácter que permiten contrastar esta hipótesis en determinados casos frente a la de un único QTL, así como modelos de caracteres múltiples para afinar la localización de la posición del QTL y discriminar entre situaciones de pleiotropía y ligamiento.

Por ejemplo, los modelos de 2 QTLs en los casos en que ambos presenten efectos aditivos o aditivos y dominantes pueden representarse de este modo:

$$M2a: y = X\beta + Zu + w_1a_1 + w_2a_2 + e$$

$$M2ad: y = X\beta + Zu + w_1a_1 + w_2d_1 + w_3a_2 + w_4d_2 + e$$

en los que se incluyen los efectos aditivo (a_2) y dominante (d_2) del segundo QTL localizado en otra posición del cromosoma, y denominamos w_3 y w_4 a los vectores de los correspondientes coeficientes c_3 y c_4 en la segunda posición considerada; la validación del segundo QTL requiere el contraste entre los modelos completo ($M2a$ o $M2ad$) y reducido ($M1a$ o $M1ad$), que representan respectivamente en este caso las hipótesis alternativa y nula.

Una de las aplicaciones de los modelos multicarácter es la posibilidad de discriminar si los efectos detectados sobre dos caracteres se deben a un solo QTL (hipótesis de pleiotropía) o a dos QTLs localizados en una misma región cromosómica (hipótesis de ligamiento). Estas dos alternativas pueden contrastarse comparando mediante la razón de verosimilitud (LR) un modelo de 2 QTLs distintos, uno para cada carácter, con un modelo de un solo QTL con efecto sobre ambos caracteres. Para la realización de estos y otros análisis se utilizaron las versiones v4.0 y anteriores del programa QXPAK (Perez-Enciso y Miształ, 2004).

3. MATERIALES Y MÉTODOS

3.6.3. Análisis de genes candidatos

Se utilizaron tres métodos para analizar la posible asociación de polimorfismos identificados en genes candidatos (GC) con los caracteres estudiados.

3.6.3.1 Modelo Animal Estándar.

El análisis permite testar la posible asociación entre los polimorfismos genotipados y los caracteres estudiados con los siguientes modelos, que implican que el polimorfismo analizado es la mutación causal o es una mutación neutral en completo desequilibrio con aquella:

$$M3a: y = X\beta + Zu + \lambda_1 g + e$$

$$M3ad: y = X\beta + Zu + \lambda_1 g + \lambda_2 h + e$$

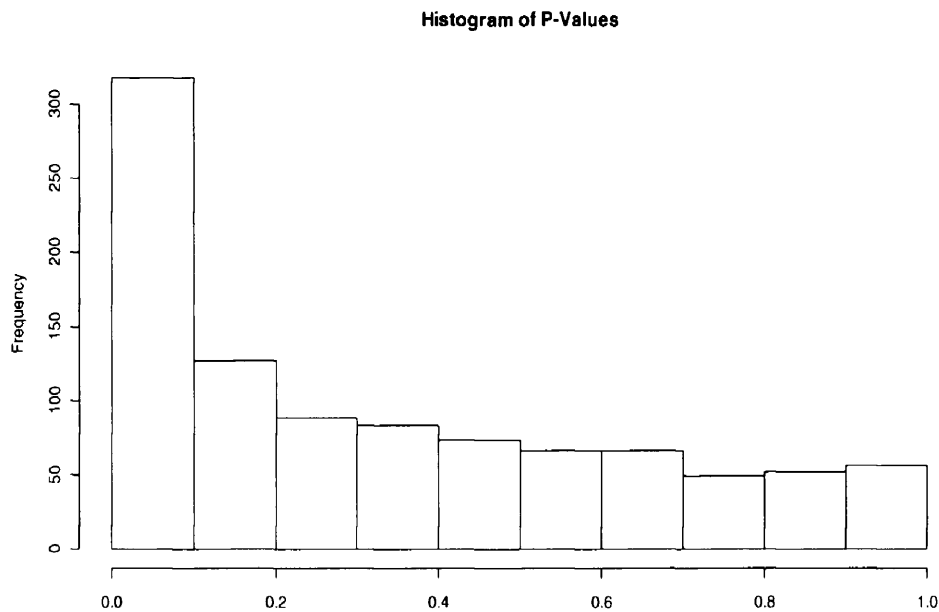
donde λ_1 es un vector $n \times 1$ que incluye una variable indicadora relacionada con el número de copias de un alelo del gen candidato y toma valores 1 o -1, si el animal es homocigoto para los alelos alternativos, y 0 si es heterocigoto; en el caso de polimorfismos bialélicos como los SNPs, λ_2 es un vector $n \times 1$ que incluye una segunda variable indicadora que toma el valor 0, si el individuo es homocigoto para cualquiera de los alelos alternativos, y 1 si es heterocigoto; g y h representan los efectos aditivo y dominante del polimorfismo analizado. La comparación, mediante el test de razón de verosimilitud, entre los modelos $M3a$ y $M0$ permite validar la significación estadística del efecto aditivo. Asimismo, la comparación análoga entre los modelos $M3ad$ y $M3a$ permite validar la significación estadística del posible efecto dominante del SNP.

Asimismo, con el modelo $M3a$ se realizó un análisis de asociación utilizando SNPs genotipados con el *PorcineSNP60 Genotyping BeadChip*. En el análisis se testaron tantas hipótesis como SNPs fueron analizados en el SSC12, esto supone por tanto la realización de múltiples tests de hipótesis con el riesgo de que muchos resulten significativos por simple azar. No es suficiente por ello el valor p de probabilidad que corresponde a un simple test de la hipótesis nula. Hay que recurrir al concepto de tasa de falso descubrimiento (FDR), que conlleva umbrales de significación más exigentes para reducir la proporción de falsos positivos. Para corregir la multiplicidad de tests se recurrió al procedimiento descrito por Storey y Tibshirani (2003), basado en la estimación de π_0 , proporción de verdaderas hipótesis nulas a partir del histograma del conjunto de valores p

3. MATERIALES Y MÉTODOS

obtenidos. A cada valor p tradicional corresponde un valor q que permite identificar como significativos los inferiores a la tasa de FDR escogida ($q < \text{FDR}$). Estos cálculos fueron realizados con el programa *Q-VALUE 1.0* facilitado por los citados autores. En los que se asumió un nivel de significación a nivel cromosómico correspondiente a un $\text{FDR} = 0.10$.

Figura 3.11. Histograma realizado por el programa *Q-VALUE*.



3.6.3.2 Modelo animal de asociación asistida por marcadores (MAAT)

Este nuevo modelo incorpora conjuntamente, además de los efectos aleatorios y sistemáticos usualmente considerados, los efectos del polimorfismo del gen candidato y del QTL para cada cM a lo largo del cromosoma (Zhao *et al.*, 2003) (Varona *et al.*, 2005). Aunque su potencia estadística es menor cuando el SNP coincide con la mutación causal, su ventaja respecto al modelo anterior de asociación, en la situación más frecuente en que el SNP es una mutación neutral, es que proporciona un menor número de falsos positivos. El caso de mayor complejidad, con aditividad y dominancia, se representaría del modo siguiente:

$$M4ad: y = X\beta + Zu + w_1a_{QTL} + w_2d_{QTL} + \lambda_1g_{SNP} + \lambda_2h_{SNP} + e$$

El contraste mediante la razón de verosimilitud (LR) entre el modelo *M4ad* y el modelo *M3ad*, que considera únicamente el efecto del SNP, permite valorar la significación

3. MATERIALES Y MÉTODOS

estadística del efecto del QTL. Alternativamente, el contraste análogo entre el modelo $M4ad$ y el modelo $M1ad$, que incluye el efecto del QTL, permite valorar la significación estadística del efecto del polimorfismo. Estos resultados permiten por tanto comparar la verosimilitud de la hipótesis de un QTL con frecuencias alternativas en las líneas parentales y la del SNP del gen candidato como subyacente al QTL detectado.

3.6.3.2 Modelos de tres alelos.

Este modelo asume que los dos alelos de un polimorfismo bialélico en un gen candidato (A y B) presentan diferente desequilibrio de ligamiento con el QTL dentro de cada línea parental. Suponiendo que el alelo A está fijado en la línea Guadyerbas y los dos alelos A y B segregan en Landrace, pueden definirse tres alelos: uno asociado con el alelo A de origen Ibérico (A_I), el segundo asociado con el alelo A de origen Landrace (A_L) y el tercero con el alelo B de origen Landrace (B). El modelo con aditividad ($M5a$) puede formularse como sigue:

$$M5a: y = X\beta + Zu + \lambda_{A_I}g_{A_I} + \lambda_{A_L}g_{A_L} + \lambda_Bg_B + e$$

dónde λ_{A_I} , λ_{A_L} y λ_B son el número de copias de los alelos A_I , A_L y B en cada individuo. Estos valores pueden calcularse combinando los genotipos del SNP y las probabilidades de origen Ibérico o Landrace, de acuerdo con el procedimiento descrito por (Braglia *et al.*, 2006). El contraste estadístico entre los modelos con tres alelos ($M5a$) y nulo ($M0$) con dos grados de libertad y entre los modelos aditivos con tres y dos alelos ($M5a$ y $M3a$) con un grado de libertad permite evaluar la significación estadística de los efectos contemplados en los mismos.

3.6.3.3 Modelo de haplotipos

Este modelo es similar al modelo de dos alelos, excepto que el término $\lambda_i g$ aparece tantas veces como número de haplotipos se quieren analizar.

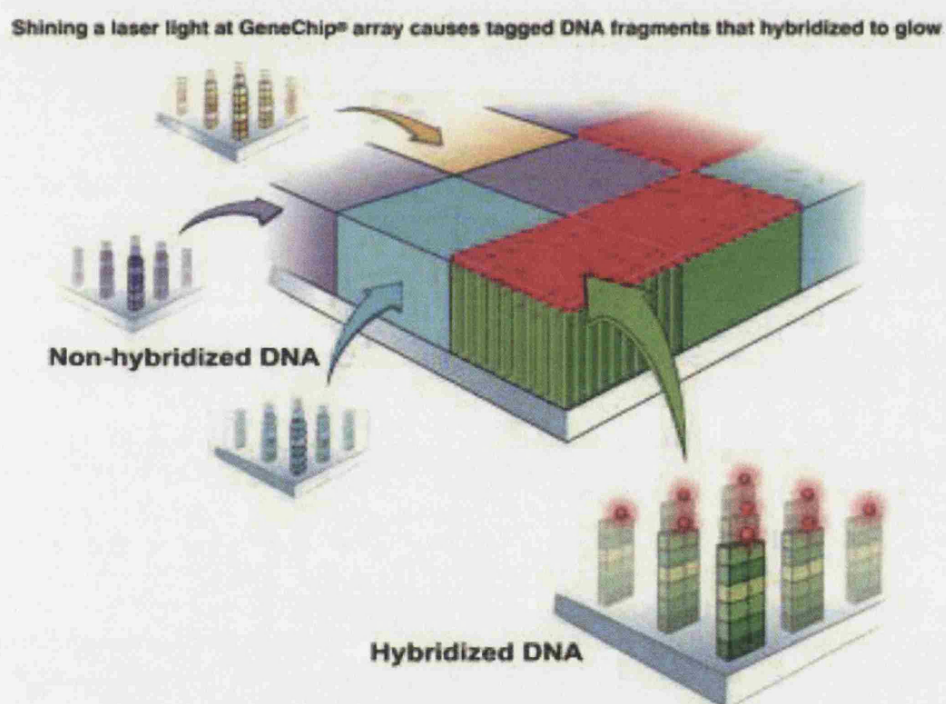
$$M6a = X\beta + Zu + \sum_{i=1}^L \lambda_i g + e$$

3. MATERIALES Y MÉTODOS

3.7. ANÁLISIS DE EXPRESIÓN CON MICROARRAYS

Se utilizaron microarrays de la plataforma comercial *Affymetrix Porcine Genechip*TM para estudiar los patrones de expresión génica en el músculo *longissimus dorsii*. El array consiste en una superficie sólida con 23937 conjuntos de sondas correspondientes a 23256 transcritos y 20201 genes. Cada conjunto de sondas, a su vez, está compuesto por 11 sondas diferentes de 25 pb que hibridan a lo largo del transcrito a estudiar.

Figura 3.12. Estructura de la plataforma *Affymetrix Porcine Genechip*TM



3.7.1. Hibridación y escaneado.

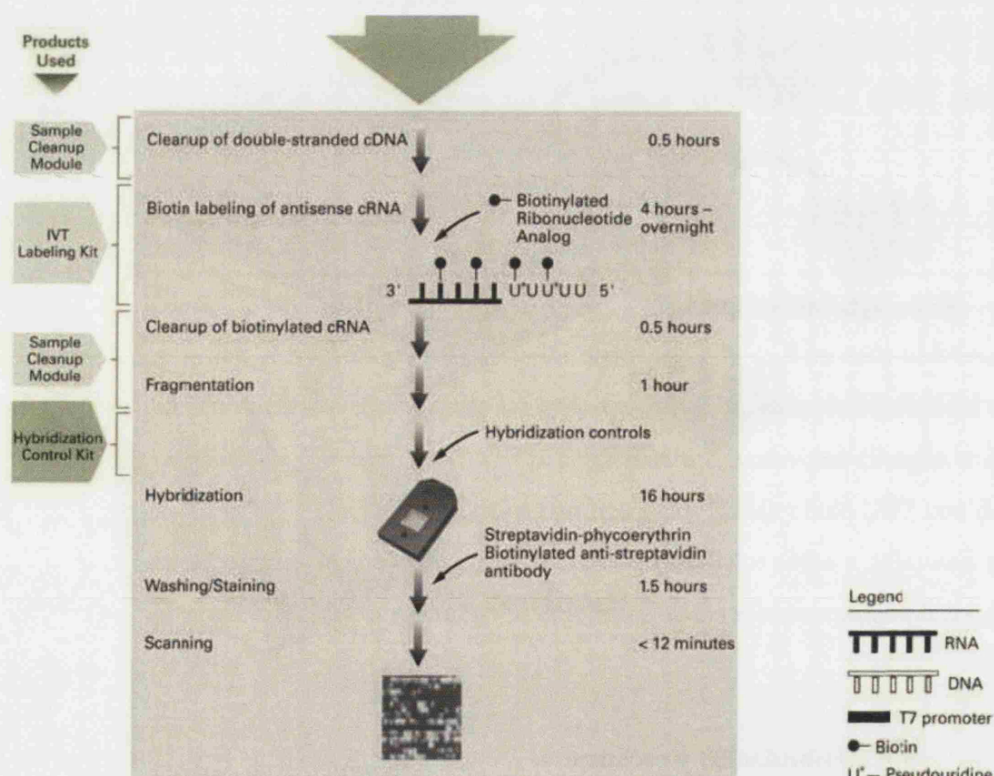
Se extrajo el ARN total de muestras de músculo *longissimus dorsii* en 100 individuos pertenecientes a la generación RC2 del cruce IBMAP y se utilizaron 7-8 µg de ARN para la hibridación en el array. En primer lugar se realizó una transcripción in vitro para sintetizar ARN complementario (ADNc), en este proceso el ADNc es fragmentado, biotinilado, hibridado y lavado sobre el array. Cada fragmento de ADNc hibrida con la correspondiente secuencia homóloga de oligonucleótidos fijada en el array. Una vez realizado este proceso el chip se escanea para evaluar la fluorescencia, de tal forma que a mayor cantidad de ADNc unido, mayor es la fluorescencia lo que indica una mayor expresión del transcrito. Las

3. MATERIALES Y MÉTODOS

extracciones de ARN y posteriores transcripciones in vitro se realizaron en el Área de Producción Animal del Centre UdL-IRTA (Lleida, España) según viene recogido en el subproyecto AGL2008-04818-C03-02.

La hibridación y escaneado de los *arrays* fue realizada en el Centre for Research in Agricultural Genomics (CRAG – Barcelona). La hibridación y escaneado se realizó mediante el protocolo estándar (Figura 3.13) de Affymetrix (Santa Clara, CA). Todo el proceso se hizo de acuerdo con las recomendaciones MIAME (Stoeckert *et al.*, 2002) desarrolladas por el *Microarray Gene expression Database Group* (<http://www.mged.org>).

Figura 3.13. Protocolo de Affymetrix para la síntesis y marcaje del ADNc



Una vez escaneados los *arrays*, los datos de fluorescencia fueron transformados en imagen pixelada (archivos.DAT) en la que cada píxel corresponde a una sonda. La media de intensidad de cada uno de los píxels es la intensidad asignada a cada sonda (archivos.CEL) y la media de las intensidades de cada sonda será el valor de la expresión de cada gen. Para realizar el análisis de estos datos se utilizó el programa *GeneChip Operation Software* (GCOS).

3.7.2. Procesamiento previo de los datos de microarrays

Antes de analizar los resultados de la hibridación se comprobó si el marcaje, hibridación y lectura del microarray había sido correcto, para evitar estimas erróneas de la expresión. Para ello se realizó un análisis visual con el programa Bioconductor (Reimers y Carey, 2006). Dentro del mismo, la rutina *Readaffy* genera una matriz de niveles de intensidad denominada *AffyBatch* a partir de los archivos .CEL. Sobre ésta se aplican tres metodologías diferentes (boxplot, histogramas sencillos y MA-plot) que se encuentran disponibles en el paquete *affydata* y que permiten examinar la intensidad de las sondas de cada uno de los *arrays*.

3.7.3. Análisis de calidad de los chips de *Affymetrix*

Antes de analizar los datos se comprobó que el ARN utilizado en la hibridación no estaba degradado mediante las siguientes funciones del programa Bioconductor:

3.7.3.1. Medida del ruido de fondo

Los valores del ruido de fondo se calcularon mediante la función *avbg* del paquete *simpleaffy*. Los valores obtenidos deben ser comparables entre los diferentes *arrays* para poder asumir que la hibridación de cada uno ha ocurrido con similar ruido de fondo, sino no sería posible realizar el análisis del conjunto de los *arrays*.

3.7.3.2. Porcentaje de llamadas

Este índice corresponde al porcentaje del conjunto de sondas denominadas como “presentes” en el *array* respecto al número total de sondas. El porcentaje debe ser similar entre los *arrays* para poder analizar los chips en su conjunto, aunque la variabilidad en el porcentaje puede indicar variabilidad biológica entre las muestras. Para la obtención del índice se ha empleado la rutina *percent.present*.

3.7.3.3. Ratio 3'/5'

Cada chip incluye varios genes que permiten controlar la calidad del ARN como, por ejemplo, los genes *GAPDH* y *β -ACTINA*. Cada gen está representado por tres sondas, una que cubre el extremo 3', otra el centro del gen y la tercera, el extremo 5'. La proporción entre la expresión de las sondas situadas en los extremos se considera una medida de la

3. MATERIALES Y MÉTODOS

calidad del ARN. Affymetrix recomienda aceptar ratios inferiores a tres. Este análisis se realizó mediante el paquete *simpleaffy*.

3.7.3.4. Calidad del ARN total

Además de la utilización del índice anterior, para poder tener una idea global de la degradación del ARN se evaluaron todas las sondas del chip. Cada transcrito está representado por 11 sondas numeradas desde el extremo 5', con lo cual, si hay degradación existirá mayor intensidad de fluorescencia en las del extremo 3' que en las del 5'.

Con la función *AffyRNAdeg* se obtuvo una representación gráfica del ratio 3'/5' de todos los *arrays* para calcular la media de la intensidad de las sondas en función de su posición en el transcrito y la pendiente de la curva que se obtiene al representar las intensidades. Una pendiente de 0.5 corresponde a un ARN de mayor calidad y una de 2 indica que está degradado.

3.7.3.5. Diagnóstico cualitativo usando un modelo PLM (a nivel de sonda)

El modelo PLM tiene en cuenta el nivel de expresión de un gen, el efecto de las sondas y el error de medición para calcular los datos de expresión normalizados con el ruido de fondo ajustado, utilizando un modelo lineal.

Mediante el paquete *affyPLM* y usando las funciones *weights*, *resids* and *sign.resids*, se pueden obtener pseudo-imágenes de los chips que simulan una imagen visual de las intensidades de sonda y de esta manera se pueden detectar posibles artefactos del *array*:

1. Imagen del signo de los residuos: hace visibles los efectos que no son apreciables en otros gráficos siendo azules los residuos negativos y rojos los positivos.
2. Imagen de pesos: usa colores topográficos, representando las zonas de mayor expresión por áreas claras y las de menor expresión con áreas oscuras.
3. Imagen de las intensidades de la sonda sin tratar (escala logarítmica).

3.7.4. Normalización

La unión inespecífica de fragmentos de ADNc a la sonda puede provocar la presencia de ruido en los chips que es un inconveniente a evitar. Para solventarlo se usan sondas adheridas al chip y diseñadas para medir las uniones inespecíficas (*mismatch probes* MM). Cada sonda que se une de manera específica a una secuencia (*perfect match* PM) tiene una sonda equivalente en la que se ha sustituido el nucleótido número 13 por otro

complementario, de tal manera que la unión deja de ser totalmente específica. Para poder obtener un valor más preciso de la señal, se resta el valor de la intensidad MM al valor de la intensidad PM. No obstante, el valor de las sondas MM suele ser mayor que el de las PM, por lo que las MM detectan tanto señal inespecífica como ruido de fondo. Otro método de normalización es el “*robust multiarray averaging*” (RMA) que corrige el ruido de fondo, normaliza por cuantiles y suma el nivel de intensidad de la sonda basándose en un modelo multiarray, aunque no tiene en cuenta la información de las MM. Para poder solventar estos problemas se ha utilizado un modelo mejorado, el GC-RMA (corrección por el contenido de GC o GeneChip-RMA) que aplica la corrección RMA pero tiene en cuenta la información de la especificidad de las sondas, generando así valores más precisos.

Para realizar la normalización se aplicó la función *gcRMA* a partir de los archivos *AffyBatch*. Además se realizó un filtrado de los datos para eliminar las sondas con más de 50% de valores perdidos.

3.7.5. Análisis de resultados

Actualmente existe una gran variedad de programas y tipos de análisis disponibles para estudiar los resultados de los *arrays*. En el presente trabajo se ha utilizado el programa GEAMM v.1.6. (Casellas *et al.*, 2008), que está escrito en lenguaje FORTRAN90 y diseñado para el análisis de modelos mixtos de expresión génica mediante inferencia bayesiana.

El programa permite realizar distintos análisis de modelos mixtos, a partir de los datos de expresión de p sondas y q arrays normalizados en logaritmo en base 2, mediante la siguiente expresión:

$$y = Xa + \sum_{i=1}^n Z_i D_i + \sum_{j=1}^m c_j \beta_j + e$$

donde y es el vector de los datos de expresión ($pq \times 1$ elementos) ordenados por sonda y por *array* e influenciados por el efecto general de cada array (a , $q \times 1$ elementos) así como por n efectos discretos (D_i) y m continuos (β_j) sobre cada sonda definidos por el usuario (ambos con dimensiones $1 \times p$). X , Z_i y c_j son las matrices (o vectores) de incidencia y e el vector de errores residuales. La expresión es desarrollada mediante inferencia Bayesiana, en la cual se asume un modelo mixto que tiene una verosimilitud bayesiana definida como distribución normal multivariante:

$$p(y | a, D_1 \dots D_n, \beta_1 \dots \beta_m, R) \sim N \left(Xa + \sum_{i=1}^n Z_i D_i + \sum_{j=1}^m c_j \beta_j, I_q \otimes R \right)$$

3. MATERIALES Y MÉTODOS

donde I_q es la matriz identidad y R la matriz de covarianzas residuales con varianzas residuales dentro de sonda y covarianzas residuales nulas entre sondas.

Este modelo considera diferentes varianzas residuales de cada sonda. Como en los análisis se utilizan test de hipótesis múltiples se aplicó un estadístico para eliminar los posibles falsos positivos que apareciesen en los test múltiples. El estadístico utilizado es la tasa de falsos descubrimientos (FDR) que se calcula a partir de la distribución de los valores p . En el contexto bayesiano los valores de la probabilidad posterior se asimilaron a los valores p frecuentistas. Fueron seleccionadas aquellas sondas con un FDR menor que 10% siguiendo el procedimiento de Benjamini y Hochberg (1995) y un valor $p < 0.001$.

3.7.6. Anotación de sondas diferencialmente expresadas

Una vez se ha obtenido la lista de las sondas diferencialmente expresadas, se procedió a anotar cada una de las mismas. Para ello, se utilizó el archivo de anotación “*Affymetrix Porcine Annotation Revision 6*” (Tsai *et al.*, 2006) que dispone de la anotación de la mayor parte de las sondas excepto un 0.39% que no están anotadas.

3.7.7. Interpretación biológica de los resultados

La interpretación biológica de los datos se llevó a cabo con la herramienta DAVID 2008 (Huang *da et al.*, 2009) la cual realiza análisis para remarcar los términos GO más relevantes de una lista de genes. Esta herramienta realiza un análisis estadístico para detectar las categorías génicas sobrerrepresentadas y las compara con el resto de sondas del *array*. La anotación funcional es reportada con un valor de p .

RESULTADOS

Caracterización de genes candidatos

4.1. Gen *Acyl-coA oxidase 1 (ACOX1)*

4.1.1. Análisis, secuenciación, búsqueda y selección de polimorfismos del ADN copia (ADNc) del gen *ACOX1*.

Se realizó una búsqueda en las bases de datos del NCBI (www.ncbi.nlm.nih.gov/) para identificar la secuencia del gen *ACOX1* en porcino. Se identificaron dos secuencias (NM_0110128 y AK_23470) en la base de datos NCBI y una en Ensembl (ENSSSCT00000018723 con un tamaño de 2291 pb). La secuencia NM_0110128 (2022 pb) está descrita como la correspondiente a la secuencia codificante del gen *ACOX1* aunque la secuencia AK_23470, que corresponde a un transcrito expresado en hígado, presenta una longitud superior de 2227pb. Por otra parte se encontraron tres variantes transcripcionales que codifican tres isoformas proteicas de este gen en humano y una secuencia transcrita en ratón. Dos de los transcritos descritos en humano sólo difieren en la secuencia del exón 3: la variante 1 (GenBank: NM_004035.6) contiene al exón 3-II y la variante 2 (NM_007292.5) contiene al exón 3-I. La tercera variante (NM_001185039.1) difiere en el extremo 5'UTR y en la secuencia codificante, y genera una proteína más pequeña en su extremo N-terminal.

Las dos secuencias de porcino depositadas en el NCBI presentan homología tanto a la variante 1 como a la variante 3 de humano, ya que su extremo 3'UTR es homólogo a la zona común a ambas variantes.

Las tres secuencias identificadas en porcino fueron alineadas junto a la de la variante 1 de humano y se procedió al diseño de oligonucleótidos que hibridasen con las regiones conservadas para amplificar el ADNc del gen en muestras de hígado. Se diseñaron siete parejas de oligonucleótidos que permiten amplificar la totalidad de la secuencia correspondiente a la isoforma 1 mediante siete fragmentos solapados (ACOX1F1/R1, ACOX1F2/R2, ACOX1F5/R5, ACOX1F6/R6, ACOX1F7/R7, ACOX1F8/R8 y ACOX1F9/R9) (Tabla 4.1).

A fin de investigar si la isoforma 2 descrita para humano también está presente en cerdo se han diseñado dos parejas de oligonucleótidos (ACOX1F3/F4 y ACOX1F4/R4) (Tabla 4.1), a partir de la secuencia NM_007292.5, que permiten amplificar el exón 3-I además de parte de los exones 2 y 4 comunes a ambas isoformas.

Tabla 4.1. Oligonucleótidos usados en la caracterización del ADNc del gen *ACOX1*

Nombre	Secuencia (5'- 3')	T _h	MgCl ₂	Longitud
ACOX1_1F	CCTCTGCTGGCTGTCACC	57°C	2.0 mM	384 pb
ACOX1_1R	GCGTTCCTGCTGTTCCCTC			
ACOX1_2F	AAAACTTTGTGCAGCCAGGAC	57°C	2.0 mM	363 pb
ACOX1_2R	CCGGATAGGGACAATGAAGG			
ACOX1_3F	ACCACGGTGAAGAAGATAAGG	50°C	2.0 mM	137 pb
ACOX1_3R	TGGCTCAGCAAGGTAGGAA			
ACOX1_4F	GGGCCTCCATTACTCCAT	54°C	2.5 mM	514 pb
ACOX1_4R	GTTACTCAGGGGCTTCACATA			
ACOX1_5F	TCATGCCCCGCCTGGAACCTT	56°C	2.0 mM	542 pb
ACOX1_5R	GCTGTAGCGGATGGCGATGGT			
ACOX1_6F	GAGGCCGCCCGGTGTCTGTC	60°C	2.0 mM	686 pb
ACOX1_6R	CACGCTGCCTCCTTGCTCTTCCT			
ACOX1_7F	CCAGCGCATCCAGCCTCAGC	56°C	2.0 mM	564 pb
ACOX1_7R	TCATGGACCTCCGTTTTGTTTAGC			
ACOX1_8F	CCTCTATGCTCTGTATGGAATC	53°C	2.0 mM	450 pb
ACOX1_8R	AAAAAGAGAAAAGGGAAAAT			
ACOX1_9F	TGCTCCGTTCCCCTCACAC	55°C	2.0 mM	458 pb
ACOX1_9R	GGCGGTCCCTCTTCAGTCC			

Para amplificar las regiones 5'UTR y 3'UTR se utilizó el kit de amplificación de ADNc SMART RACE (Clontech Laboratories) y los oligonucleótidos ACOX1R, en el caso del extremo 5'UTR, y ACOX9F, en el caso del extremo 3'UTR. En el caso del extremo 5'UTR solamente se han podido aumentar la longitud de la secuencia conocida en 9 pb y en el caso del extremo 3'UTR no se consiguió amplificar ningún fragmento que contuviese un mayor número de pares de bases.

El ensamblaje de los fragmentos amplificados ha dado lugar a dos secuencias con una longitud de 2457 pb. Estas dos variantes difieren únicamente en la secuencia del exón 3, aunque ambas codifican para sendas proteínas compuestas por 661 aminoácidos. Una de ellas presenta un 74.8% de homología con la variante 1 humana y un 68.2% con la de ratón, mientras la otra presenta un 74.8% de homología con la variante 2 humana. Se amplificó la totalidad de la secuencia a partir del ADNc de hígado de dos animales Ibéricos (línea Guaduerbas) y dos animales Landrace. Mediante el ensamblaje y alineamiento de las

secuencias se identificaron cinco polimorfismos de un solo nucleótido (SNP), cuatro en la secuencia codificante y uno en la región 3'UTR (Tabla 4.2). Tres de los cuatro SNPs detectados en la secuencia codificante dan lugar a cambio de aminoácido.

Tabla 4.2. Polimorfismos detectados en el ADNc del gen ACOX1

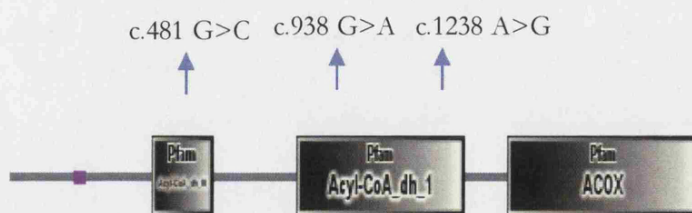
Gen	Posición	Exon	Cambio Aminoacídico
ACOX1	c.481G>C	4	Glu 162 Gln
	c.938G>A	7	Arg 313 Lys
	c.1101A>G	8	Sinónimo, Leu
	c.1238A>G	9	Asn 413 Ser
	c.*21G>A	3'UTR	-

Para evaluar el efecto potencial de las mutaciones se utilizaron herramientas *in silico* disponibles en internet de manera gratuita.

De acuerdo con las predicciones del programa SIFT, los polimorfismos c.938 G>A y c.1238 A>G producen cambios tolerados mientras el polimorfismo c.481 G>C genera un cambio de aminoácido no tolerado.

Mediante el programa SMART se determinó en qué dominios proteicos se encuentran los polimorfismos detectados. Estos dominios se representan en la figura 4.1.

Figura 4.1: Dominios proteicos de la proteína ACOX1 porcina



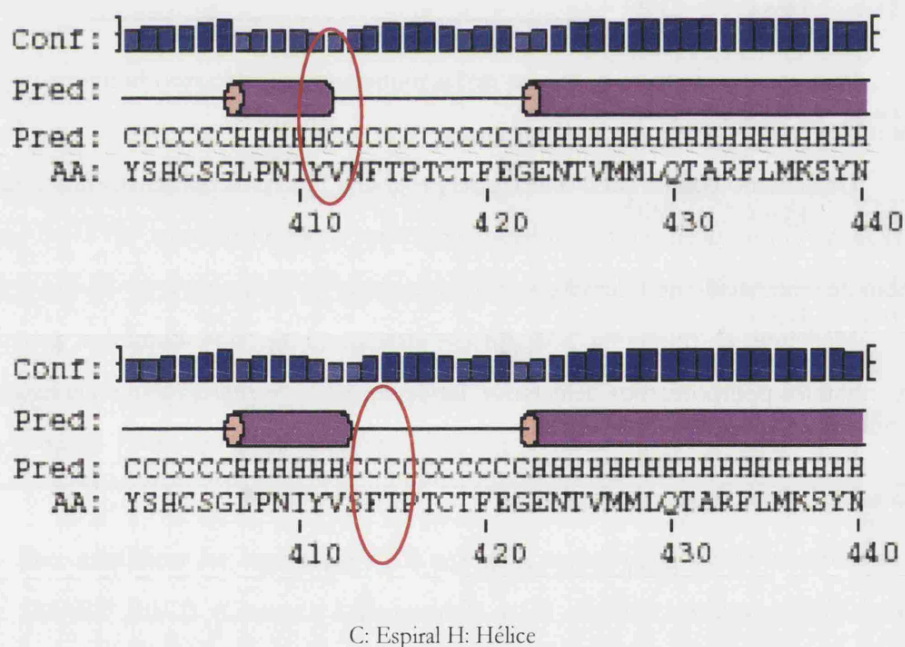
(■) representa una zona de baja complejidad proteica, los dos siguientes (■) son dominios típicos de deshidrogenasas que son proteínas flavoproteínas que catalizan la deshidrogenación de acyl-CoA thioesters a trans 2,3-enoyl CoA. El último dominio, *Pfam ACOX* es característico de la proteína ACOX1 y es el que le confiere la función

El dominio *Pfam ACOX* es característico de la proteína ACOX1 y es el que le confiere la función, por lo tanto los cambios aminoacídicos en estos dominios podrían alterar la función de la proteína e incluso dar lugar a una proteína no funcional. El cambio

aminoácido producido por c.481 G>C afecta al dominio *Pfam Acyl-CoA_dh_M* y los cambios generados por los polimorfismos c.938 G>A y c.1238 A>G se localizan en el dominio *Pfam Acyl-CoA_dh_1*.

La herramienta PSIPRED permitió detectar un cambio en la estructura secundaria de la proteína producido por el polimorfismo c.1238 A>G (Figura 4.2). Este software predice que la sustitución del aminoácido Asparagina por Serina alteraría la estructura conformacional de la proteína incrementando el tamaño de una de las hélices.

Figura 4.2: Simulación *in silico* con el programa PSIPRED de la estructura secundaria de la proteína ACOX1 correspondiente a la región en donde se ha detectado el polimorfismo c.1238 A>G.



Por último mediante el uso de las herramientas Targetscan y Pictar se analizó si el polimorfismo c.*21 G>A localizado en la región 3'UTR podría afectar a dianas de microARNs. No se detectó ninguna diana de microARN en la región que contiene este polimorfismo aunque esto no tiene porque ser determinante debido a que las bases de datos están principalmente basadas en los genomas de humano y ratón.

Basándonos en estos resultados, los polimorfismos c.481G>C, c.1238A>G y c.*21G>A fueron seleccionados para ser genotipados en los parentales del cruce. A pesar de que el polimorfismo c.938G>A produce un cambio de aminoácido pero no se ha

podido genotipar porque se encuentra en el extremo de un exón y no se dispone de la secuencia intrónica adyacente.

4.1.2. Genotipado de polimorfismos del gen *ACOX1*

Para el genotipado de los polimorfismos seleccionados se diseñaron tres protocolos de pirosecuenciación (Tabla 4.3).

Tabla 4.3. Oligonucleótidos utilizados en el genotipado de los polimorfismos c.481 G>C, c.1238 A>G y c.*21 G>A.

Nombre	Secuencia (5'- 3')	T _b	MgCl ₂	Longitud
Pyr-c.481F	Biot-CTTGGAACACAGCCACTTA			
Pyr-c.481R	GTCACAGTAGGGCTATTGAGAATG	54°C	2.0 mM	66 pb
Pyr-c.481S	GCTATTGAGAATGAACTCCT			
Pyr-c.1238F	Biot-CGGGCACGGCTATTCTCACT			
Pyr-c.1238R	AAGGTGCAGGTGGGGTGAA	56°C	2.0 mM	69 pb
Pyr-c.1238S	GCAGGTTGGGGTGAA			
Pyr-c.*21F	Biot-CCTGACGACGAGTTCAGTCTG			
Pyr-c.*21R	AGGTTCCACGCAGATTTCTGT	60°C	2.0 mM	83 pb
Pyr-c.*21S	CGGAGCACAACCTGCTT			

En una primera fase se genotiparon los tres polimorfismos en los parentales del IBMAP (F0). Se observó que los tres polimorfismos están cosegregando lo que ha permitido definir dos haplotipos:

- Haplotipo 1 (H1): c.[481G; 1238A; *21G]
- Haplotipo 2 (H2): c.[481C; 1238G; *21A]

Los haplotipos fueron confirmados en las madres Landrace de los dos retrocruces (RC1 y RC2). En el resto de los individuos del cruce IBMAPse genotipo únicamente uno de los polimorfismos (c.481G>C) y se ha inferido el genotipo de los otros dos. Por lo tanto los genotipos GG, GC y CC serían equivalentes a las combinaciones haplotípicas H1H1, H1H2 y H2H2 respectivamente. Las frecuencias de los haplotipos en el total de los

individuos del cruce se representan en la tabla 4.4, observándose frecuencias haplotípicas extremas, con H1 mayoritaria frente a la de H2 en todos los casos, excepto en los parentales ibéricos.

Tabla 4.4. Frecuencias haplotípicas de H1: c.[481G; 1238A; *21G] y H2: c.[481C; 1238G; *21A] en los individuos parentales, F2, F3, RC1 y RC2.

Generación	Número	Frec. H1	Frec. H2
F0	Padres Ib (3)	0.18	0.82
	Madres Ld (30)	0.99	0.01
F2	Individuos (390)	0.88	0.12
Parentales F3	Padres F2 (3)	0.83	0.17
	Madres F2 (15)	0.93	0.07
F3	Hijos (52)	0.92	0.09
Parentales RC1	Padres F2 (4)	0.75	0.25
	Madres Ld (22)	1	0
RC1	Individuos (79)	0.86	0.14
Parentales RC2	Padres F1 (5)	1	0
	Madres Ld (25)	0.98	0.02
RC2	Individuos (142)	0.99	0.01
F3+RC1+RC2	Individuos (273)	0.94	0.06

4.2. Gen *ATP citrate lyase (ACLY)*

4.2.1. Análisis, secuenciación, búsqueda y selección de polimorfismos del ADNc del gen *ACLY*

Se realizó una búsqueda en las bases de datos anteriormente mencionadas de las secuencias del gen *ACLY* en porcino, humano y ratón. Se identificó una secuencia en

porcino (NM_001105302) con 4378 pb, dos en humano y una en ratón. Las dos variantes descritas en humano codifican dos isoformas que difieren en la presencia o ausencia del exón 14. La secuencia más larga, NM_001096.2, codifica para la isoforma proteica 1, que incluye el exón 14 y la secuencia NM_198830.1 codifica para la isoforma proteica 2 que carece de dicho exón. El alineamiento de estas secuencias indica homología entre los transcritos descritos en porcino, ratón y la isoforma 2 del gen *ACLY* humano.

Se diseñaron 11 parejas de cebadores (Tabla 4.5) que hibridan desde el exón 2 hasta el exón 27 de la secuencia NM_001105302. Los fragmentos fueron amplificados a partir

Tabla 4.5. Oligonucleótidos usados en la caracterización del ADNc del gen *ACLY*

Nombre	Secuencia (5'- 3')	T _b	MgCl ₂	Longitud
ACLY1F	CCGCCCAAGCCGGAACAAAG	61°C	2.5 mM	225 pb
ACLY1R	TTCTGGATGGCCGAGGTGGTG			
ACLY2F	GTCGGCCAAGGCGATTTTGGAGC	61°C	2.0 mM	699 pb
ACLY2R	CCCCACTTCACTTTGCAGATGTAG			
ACLY3F	TGCCCCCGAAGACAAGAAAGAAAT	56°C	2.0 mM	654 pb
ACLY3R	ACCCGTAAGCCCTCCTGATAGTTG			
ACLY4F	CGAAGAGGCGGCCCAACTAT	60°C	2.5 mM	357/387 pb
ACLY4R	GGTCTGCATGCCCCACACGAT			
ACLY5F	CCTCTTCAGCCGCCACACCAA	61°C	2.5 mM	529 pb
ACLY5R	GGACATGCCGCCCGAACG			
ACLY6F	CCGGCGGGATGCTGGATAACA	59°C	2.0 mM	519 pb
ACLY6R	GCACGATGACCCCTCTGGACACA			
ACLY7F	GCCGGGGCATCAAGGAAG	58°C	2.5 mM	515 pb
ACLY7R	ATCGGCCGTCACCATCAGG			
ACLY8F	GGCATCGGCGGGGTCCTC	60°C	2.0 mM	426 pb
ACLY8R	TTCCAGCGCGTAGTCAAGCAGA			
ACLY9F	TAGACAACCCAGACATGAGAGTGC	56°C	2.0 mM	561 pb
ACLY9R	TGTTAGCCTGTGGGTATTGTTTTTC			
ACLY10F	GCTGGCACTGGCGCTTACTTA	54°C	2.0 mM	529 pb
ACLY10R	TCCCGACCCCTGACACA			
ACLY11F	CGAATGTCTTGTTCATTG	56°C	2.5 mM	414 pb
ACLY11R	TGCAGAGGAGAGACTGAGGTG			

del ADNc. Asimismo se diseñó una pareja de oligonucleótidos, ACLY4F/4R, en los exones 11 y 15 que, mediante la amplificación del exón 14, permite verificar en hígado la presencia de las dos variantes descritas en humano. El análisis de los productos amplificados mediante electroforesis en gel de agarosa al 3% ha permitido identificar dos bandas. Se secuenciaron estos fragmentos y se compararon las secuencias resultantes con 357 pb y 387 pb, respectivamente con la secuencia de humano. Se verificó que el fragmento adicional de 30 pb presenta homología con el exón 14 de la variante 1.

Se caracterizaron dos transcritos en dos animales Guadyerbas y dos Landrace, uno de 4116 pb homólogo a la variante 1 de humano y otro de 4086 pb homólogo a la variante 2. En ambos casos la homología con las secuencias de humano es de un 87.5%. Se detectaron dos polimorfismos (Tabla 4.6) uno silente y otro en la región 3'UTR.

Tabla 4.6. Polimorfismos detectados en el ADNc del gen *ACLY*

Gen	Posición	Exón	Cambio Aminoácido
<i>ACLY</i>	c.2790C>T	24	Sinónimo, Pro
	c.*523T>C	3'UTR	-

Se analizó si el polimorfismo detectado en la región 3'UTR podría afectar a alguna diana de microARNs pero ninguna de las herramientas utilizadas y bases de datos consultadas generó ningún resultado.

4.2.2. Análisis, secuenciación y búsqueda de polimorfismos en la región promotora del gen *ACLY*

Ren *et al.* (2008) describieron la secuencia de la región promotora del gen *ACLY* (Nº de acceso GenBank: EU073663) en porcino. Se diseñaron dos parejas de oligonucleótidos para amplificar dos fragmentos solapados (Tabla 4.7) permitiendo el conocimiento de 755 pb de los 870 descritos. Estos fragmentos se amplificaron en ADN genómico de los parentales Ibéricos y Landrace del cruce. Sin embargo, después del ensamblaje y alineamiento de las secuencias ningún polimorfismo fue detectado.

Tabla 4.7. Oligonucleótidos usados en la secuenciación de la región promotora del gen *ACLY*.

Nombre	Secuencia (5'- 3')	T _h	MgCl ₂	Longitud
ACLYPROM1F	AGGGGGAGGCAGCCATTCATTCA	63°C	2.0 mM	694 pb
ACLYPROM1R	CCTCGGGTCCGGGGTCCAG			
ACLYPROM2F	ACACCGCAGCCCGCCACATC	63°C	2.0 mM	450 pb
ACLYPROM2R	GCCCAACTCGCCGCCTACCTTC			

4.2.3. Genotipado de polimorfismos del gen *ACLY*.

Se genotiparon los dos polimorfismos detectados mediante dos protocolos de pirosecuenciación (Tabla 4.8).

Tabla 4.8. Oligonucleótidos utilizados en el genotipado de los polimorfismos c.2790 C>T y c.*523 T>C. del gen *ACLY*

Nombre	Secuencia (5'- 3')	T _h	MgCl ₂	Longitud
Pyr-c.2790F	GTTCGGGGGTGCCCTGGA	64°C	2.0 mM	65 pb
Pyr-c.2790R	Biot-GGATGATGCCGCTGTCAAAA			
Pyr.c.2790S	GTTCGGGGGTGCCCT			
Pyr-c.*523F	Biot-CCCTTGGAGGAAAGGAATGA	60°C	2.5 mM	110 pb
Pyr-c.*523R	TCAGGTTATATGAGGAGGAAGAAA			
Pyr.c.*523S	GTTATATGAGGAGGAAGAAA			

En la tabla 4.9 se representan las frecuencias alélicas de ambos SNPs. En una primera etapa, ambos polimorfismos fueron genotipados en la F0 para comprobar su segregación en el cruce. Se observó que el polimorfismo c.2790C>T segrega con frecuencias extremas, mientras que el polimorfismo c.*523 T>C segrega a frecuencias intermedias. Ambos polimorfismos fueron genotipados en el resto de generaciones del ILMAP exceptuando el RC2 que se genotipó solamente para el SNP c.*523 T>C por ser más informativo.

Tabla 4.9. Frecuencias alélicas de los polimorfismos c.2790 C>T y c.*523 T>C del gen *ACLY* en el cruce IBMAP

	Generación	Número	Alelo1	Alelo 2
c.2970T>C	F0	Padres Ib (3)	1	0
		Madres Ld (30)	0.68	0.32
	F2	Individuos (388)	0.81	0.19
	Parentales F3	Padres F2 (3)	0.83	0.17
		Madres F2 (15)	0.87	0.13
	F3	Individuos (55)	0.87	0.13
	Parentales RC1	Padres F2 (4)	0.88	0.12
		Madres Ld (22)	0.68	0.32
	RC1	Individuos (78)	0.78	0.22
	F3+RC1	Individuos (133)	0.82	0.18
c.*523 T>C	F0	Padres (3)	0.83	0.17
		Madres (30)	0.42	0.58
	F2	Individuos (387)	0.43	0.57
	Parentales F3	Padres (3)	0.5	0.5
		Madres (15)	0.37	0.63
	F3	Individuos (55)	0.39	0.61
	Parentales RC1	Padres (4)	0.63	0.37
		Madres (22)	0.55	0.45
	RC1	Individuos (79)	0.61	0.39
	Parentales RC2	Padres (5)	0.40	0.60
		Madres (25)	0.36	0.64
	RC2	Individuos (137)	0.42	0.58
	F3+RC1+RC2	Individuos (271)	0.47	0.53

4.3. Gen *Sterol regulatory element binding factor 1 (SREBF1)*4.3.1. Análisis, secuenciación, búsqueda y selección de polimorfismos del ADNc del gen *SREBF1*

Se utilizó la secuencia NM_214157, descrita como secuencia completa del mRNA del gen *SREBF1* en *Sus scrofa*. Esta secuencia contiene 3830 pb, estructurados en 20 exones que codifican para una proteína de 1151 aminoácidos y es homóloga a las secuencias de este gen descritas en humano y ratón respectivamente. Se diseñaron siete parejas de oligonucleótidos que permiten la amplificación de siete fragmentos solapados, desde el exón 1 hasta el exón 20 (Tabla 4.10). Tras ensamblar los siete fragmentos se obtuvo una secuencia de 3619 pb. Esta secuencia presenta un 86% de homología con la secuencia descrita en humano.

Tabla 4.10. Oligonucleótidos utilizados en la caracterización del ADNc del gen *SREBF1*

Nombre	Secuencia (5'- 3')	T _b	MgCl ₂	Longitud
SREBF1_1F	CGCCATGGACGAGCCGCCCTTCA	65°C	2.0 mM	664 pb
SREBF1_1R	GGGGGTGGCCGTGGCTGTCAACA			
SREBF1_2F	TCTCCGTGCACACCCAGGTTCAAA	63°C	2.0 mM	691 pb
SREBF1_2R	GGGGGCGGGCTAAGGGTGTCTAC			
SREBF1_3F	TTGCGCAAGGCCATCGACTACAT	63°C	2.0 mM	629 pb
SREBF1_3R	CGCGTGACCGGCTCTCCATAG			
SREBF1_4F	GCCCTGTGCGTGCTCGTCTTCCTC	60°C	2.0 mM	876 pb
SREBF1_4R	GCACACGGCCCAGTCCCCATCCAC			
SREBF1_5F	GCTGCCCTGCGAGTCAAGACC	60°C	2.0 mM	691 pb
SREBF1_5R	GGTACCCACTGGCCTTCTCACAC			
SREBF1_6F	CCGCGGGAAGGCAGAGTC	60°C	2.0 mM	519 pb
SREBF1_6R	CCGGGCGGCAGGTAGCAG			
SREBF1_7F	CTGCCATGCGGAGGGTGTTC	60°C	2.0 mM	527 pb
SREBF1_7R	CGGGGTGCGGGCTGGAC			

Se amplificó el ADNc extraído de muestras de músculo *longissimus dorsi* de dos animales Guadyerbas, dos Landrace y dos individuos pertenecientes al RC2. Tras el ensamblaje y alineamiento de las secuencias se detectaron seis polimorfismos (Tabla 4.11)

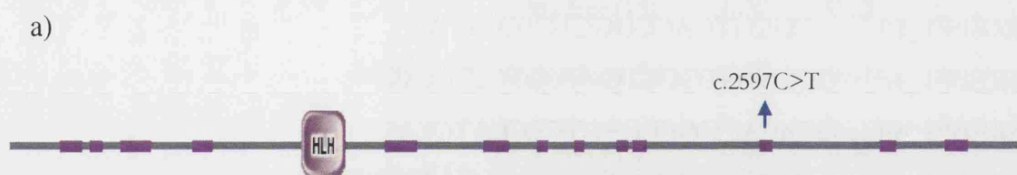
en la región codificante del gen, cinco de ellos sinónimos y uno, el SNP c.2597C>T, produce un cambio de Treonina a Metionina en la posición 866 de la proteína.

Tabla 4.11. Polimorfismos detectados en el ADNc del gen *SREBF1*

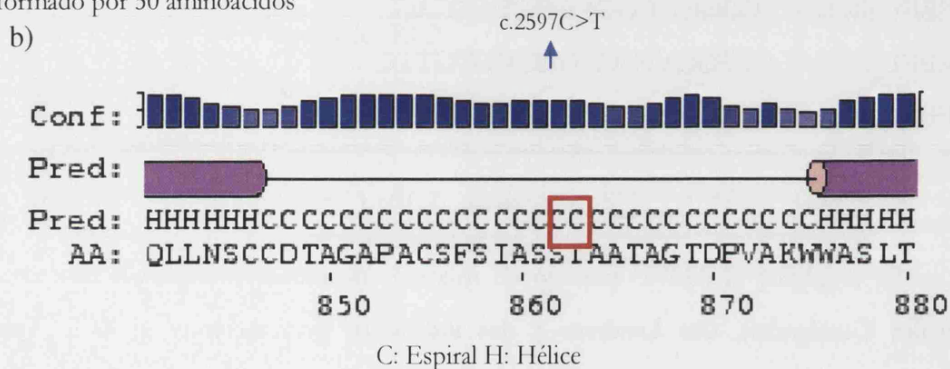
Gen	Posición	Exón	Cambio Aminoacídico
<i>SREBF1</i>	c.111C>T	2	Sinónimo, Asn
	c.219G>A	2	Sinónimo, Pro
	c.996C>T	5	Sinónimo, Ile
	c.1023T>C	5	Sinónimo, Asn
	c.1035C>T	5	Sinónimo, Ile
	c.2597C>T	15	Thr 866 Met
	c.2613G>A	16	Sinónimo, Pro

Se realizó un análisis *in-silico* para investigar la posible relevancia de estos polimorfismos mediante las herramientas disponibles en internet.

Figura 4.3.a) Predicción de dominios estructurales de la proteína SREBF1 detectados con la aplicación SMART. b) Simulación *in-silico* generada con el software PSIPRED de la estructura secundaria de la proteína SREBF1 para la región que contiene el aminoácido alterado por el SNP c.2597 A>G.



Los rectángulos rosas representan dominios de baja complejidad y el cuadrado más grande con las siglas HLH indica la existencia de un dominio hélice-bucle-hélice, este dominio está formado por 50 aminoácidos



De acuerdo con las predicciones de la herramienta SIFT, el polimorfismo c.2597C>T produce un cambio tolerado por la estructura de la proteína que, según el modelo generado por las herramientas SMART y PSIPRED está localizado en una región de baja complejidad (Figura 4.3). Las proteínas que contienen los dominios hélice-bucle-hélice, como parece ser el caso de la ACOX, actúan como factores de transcripción eucariotas con una influencia determinante en una gran variedad de rutas de desarrollo y están muy conservados entre las diferentes especies.

Según las predicciones del “RNA Secondary Structure Prediction” los SNPs c.996C>T, c.1023A>G y c.2597C>T producirían cambios estructurales en la conformación secundaria del RNA, por lo que se decidió llevar cabo su genotipado en el material IBMAP.

4.3.2. Genotipado de polimorfismos del gen *SREBF1*.

Dos protocolos de amplificación y secuenciación fueron diseñados para verificar la segregación de los polimorfismos en la F0 (Tabla 4.12). La pareja de oligonucleótidos SREBF1Ex5F/SREBF1Ex5R amplifica un fragmento del exón 5, de 162 pb, que contiene los SNPs c.996 C>T y c.1023 T>C. Los fragmentos amplificados fueron alineados y se detectó un nuevo SNP, c.1035 C>T, con frecuencias extremas en los individuos parentales (Tabla 4.15)

Tabla 4.12. Oligonucleótidos empleados en la secuenciación del exón 5 del gen *SREBF1*

Nombre	Secuencia (5'- 3')	T _b	MgCl ₂	Longitud
SREBF1Ex5F	ACAAGCTGCCCATCAACCG	66°C	2.0 mM	162 pb
SREBF1Ex5R	TGCCCAACCAGGTCCTT			

El SNP c.2597C>T se encuentra a 5 pb del extremo del exón 15 lo que imposibilita el diseño de cebadores sin el conocimiento de la secuencia intrónica adyacente. Se realizó una búsqueda de secuencias genómicas de humano y ratón que contuviesen el intrón 15-16 y se alinearon con las secuencias traza de porcino depositadas en las bases de datos para detectar alguna secuencia homóloga a las de humano (nº acceso GenBank: NT_010718.16) y ratón (nº acceso GenBank: NT_096135.5|Mm11_95772_37). Se detectó una secuencia traza, dpcxa0_023376.z1, partir de la cual se realizó el diseño de oligonucleótidos (Tabla 4.1.3) que permitieron amplificar y secuenciar un fragmento de 318 pb en la F0 del cruce.

Se verificó que el SNP c.2597C>T está segregando en la generación parental del IBMAP con frecuencias intermedias (Tabla 4.16).

Tabla 4.13. Oligonucleótidos empleados en la secuenciación del fragmento del ADN genómico que contiene parte de los exones 15 y 16 del gen *SREBF1*

Nombre	Secuencia (5'- 3')	T _b	MgCl ₂	Long
SREBF1Ex15_16F	CGCTCGGCCGCCTCCTCGTC	64°C	1.5 mM	318 pb
SREBF1Ex15_16R	TGTCCCAATGGCCACTCCTCTCCT			

Comprobada la segregación de estos cuatro polimorfismos, se diseñaron protocolos para su genotipado mediante pirosecuenciación. Como los polimorfismos c.996C>T, c.1023 T>C y c.1035 C>T están todos localizados en el exón 5 con una distancia de 162pb entre el primero y el último, se amplificó esa región con la pareja de oligonucleótidos SREBF1Ex5F/SREBF1Ex5R empleada en la secuenciación (Tabla 4.12), tras marcar el oligo reverse con biotina en su extremo 5'. Debido a que el SNP c.996C>T se encuentra a 27 pb de los otros dos SNPs, se diseñaron dos oligonucleótidos (Pyr- SNP1S y Pyr- SNP2_3S) de secuenciación para emplear en sendas reacciones de pirosecuenciación. En una reacción se genotipó el SNP c.996C>T y en la segunda los SNPs c.1023T>C y c.1035C>T. Además, se diseñó un tercer protocolo de pirosecuenciación para genotipar el polimorfismo c.2597C>T (Tabla 4.14). Los cuatro polimorfismos fueron genotipados en las generaciones F1 y F2 del cruce IBMAP y las frecuencias alélicas obtenidas se representan en la tabla 4.15.

Tabla 4.14. Oligonucleótidos empleados en el genotipado por pirosecuenciación de los polimorfismos c.996C>T, c.1023A>G, c.1035 C>T y c.2597C>T.

Nombre	Secuencia (5'- 3')	T _b	MgCl ₂	Longitud
SREBF1Ex5F	ACAAGCTGCCCATCAACCG	66°C	2.5 mM	162 pb
B-SREBF1Ex5R	Biot-TGCCCACCACCAGGTCCTT			
Pyr- SNP1S	AAGCGGACGGCTCAC			
Pyr-SNP2_3S	CTACCGCTCCTCCAT			
Pyr-2597F	Biot-GCTCCTGCCTGCAGCTTCT	60°C	1.5 mM	110 pb
Pyr-2597R	AAACCCAGCCACACTGCACTC			
Pyr-2597S	CCTGGGGGCTCACCG			

Tabla 4.15. Frecuencias alélicas de los polimorfismos c.996C>T, c.1023A>G, c.1035 C>T y c.2597C>T en los individuos parentales y F2.

Polimorfismo	Generación	Número	Alelo1	Alelo 2
c.996C>T	F0	Padres Ib (3)	1	0
		Madres Ld (30)	0.80	0.20
	F2	Individuos (386)	0.77	0.23
c.1023T>C	F0	Padres (3)	1	0
		Madres (30)	0.52	0.48
	F2	Individuos (388)	0.69	0.31
c.1035C>T	F0	Padres (3)	1	0
		Madres (30)	0.80	0.20
	F2	Individuos (385)	0.93	0.07
c.2597C>T	F0	Padres (3)	0.17	0.83
		Madres (30)	1	0
	F2	Individuos (378)	0.54	0.46

Los polimorfismos c.1023T>C y c.2597C>T fueron los más informativos y, por ello se genotiparon en las restantes generaciones (F3, RC1 y RC2) del cruce IBMAP (Tabla 4.16).

Tabla 4.16. Frecuencias alélicas de los polimorfismos c.1023T>C y c.2597C>T en los individuos F3, RC1 y RC2.

	Generación	Número	Alelo1	Alelo 2
c.1023T>C	Parentales F3	Padres (3)	0.67	0.33
		Madres (15)	0.63	0.37
	F3	Individuos (53)	0.68	0.37
	Parentales RC1	Padres (4)	0.75	0.25
		Madres (22)	0.70	0.30
	RC1	Individuos (77)	0.83	0.17
	Parentales RC2	Padres (5)	0.90	0.10
		Madres (25)	0.74	0.26
	RC2	Individuos (136)	0.90	0.20
	F3+RC1+RC2	Individuos (266)	0.77	0.23
c.2597C>T	Parentales F3	Padres (3)	0.67	0.33
		Madres (15)	0.67	0.33
	F3	Individuos (54)	0.72	0.28
	Parentales RC1	Padres (4)	0.5	0.5
		Madres (22)	0	1
	RC1	Individuos (76)	0.75	0.25
	Parentales RC2	Padres (5)	0.5	0.5
		Madres (25)	0	1
	RC2	Individuos (139)	0.75	0.25
	F3+RC1+RC2	Individuos (269)	0.74	0.26

4.4. Gen *Acetyl-CoenzymeA carboxilase alpha (ACACA)*

4.4.1. Búsqueda de secuencia promotora, secuenciación y detección de polimorfismos.

Aunque no se dispone de información referente a porcino, se sabe que en otras especies de mamíferos la transcripción del gen *ACACA* está controlada por tres promotores (PI, PII y PIII). Este trabajo se ha centrado en la búsqueda de polimorfismos en la región correspondiente a PI. Tras una búsqueda en las bases de datos de posibles secuencias promotoras de otras especies se utilizaron las secuencias correspondientes a PI identificadas en humano (n° de acceso GenBank: AY315616) por (Mao *et al.*, 2003) y bovino (n° de acceso GenBank: AJ276223), por (Mao *et al.*, 2001) para buscar secuencias homólogas en porcino. No se detectó ninguna secuencia homóloga a la región PI de humano pero sí a la de bovino. La secuencia encontrada corresponde al clon CH242-27L18 (n° de acceso GenBank: FJ263680), que contiene la secuencia del gen *ACACA* en porcino. Se ha encontrado una homología de un 82% en 442 nt, siendo 223 nt localizados en una región intrónica entre los exones 1 y “1c-like” y 219 en el exón “1c-like”. Se introdujo este fragmento de 442 pb pares de bases en los programas MATINSPECTOR y TFSEARCH habiéndose detectado numerosos motivos de unión a factores de transcripción. Se diseñó una pareja de oligonucleótidos para secuenciar esta región en los parentales del cruce (Tabla 4.17) y detectar posibles polimorfismos.

Tabla 4.17. Oligonucleótidos utilizados en la secuenciación del posible PI del gen porcino *ACACA*.

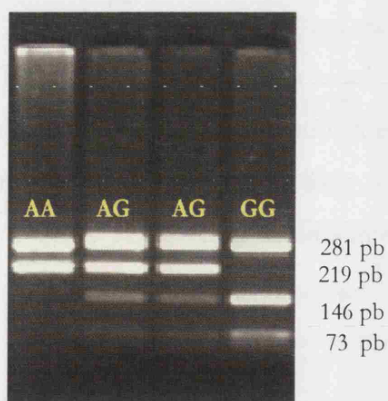
Nombre	Secuencia (5'- 3')	T _b	MgCl ₂	Longitud
ACACAPROMF	CCTGGGATGAATGGAGACACTG	58°C	2.5 mM	500 pb
ACACAPROM	GCAAATACCCTGAGAAAGCCTAAG			

Las secuencias fueron editadas y alineadas y se detectó un polimorfismo, *ACACA* g.48001 A>G, localizado a 178 pb del extremo 5' del fragmento homólogo a la región PI bovina.

4.4.2. Genotipado del polimorfismo *ACACA* g.48001 A>G

El polimorfismo se encuentra en una zona con un alto contenido en GC y su genotipado por pirosecuenciación fue inviable. Se diseñó entonces un protocolo de genotipado mediante PCR-RFLP. La reacción de amplificación se realizó con la misma pareja de oligonucleótidos que se había empleado en la secuenciación. El polimorfismo *ACACA* g.48001 A>G genera una diana polimórfica para el enzima *MspI*, además de otra diana monomórfica que está presente en el mismo fragmento. De tal manera, que los individuos homocigotos AA, presentan dos fragmentos digeridos (281 y 219 pb), los homocigotos GG tres fragmentos (281, 146 y 73 pb) y los heterocigotos cuatro. Las reacciones de PCR-RFLP se visualizaron mediante electroforesis en gel de agarosa al 2.5% (Figura 4.5).

Figura 4.5. Patrón electroforético del polimorfismo *ACACA* g.48001A>G PCR-RFLP *MspI*.



Mediante esta técnica se genotiparon todos los individuos del cruce experimental Ibérico x Landrace detectándose una segregación con frecuencias intermedias en las generaciones F2, F3 y ambos retrocruces (Tabla 4.18).

Tabla 4.18. Frecuencias alélicas del polimorfismo *ACACA* g.48001A>G en el cruce IBMAP

Polimorfismo	Generación	Número	Alelo 1	Alelo 2
g.48001 A>G	F0	Padres (3)	0.17	0.83
		Madres (30)	0.60	0.40
	F2	Individuos (391)	0.59	0.41
	Parentales F3	Padres (3)	0.33	0.67
		Madres (15)	0.57	0.43
	F3	Individuos (55)	0.45	0.55
	Parentales RC1	Padres (4)	0.38	0.62
		Madres (22)	0.55	0.45
	RC1	Individuos (78)	0.46	0.54
	Parentales RC2	Padres (5)	0.70	0.30
		Madres (25)	0.34	0.66
	RC2	Individuos (145)	0.52	0.48
	F3+RC1+RC2	Individuos (278)	0.49	0.51

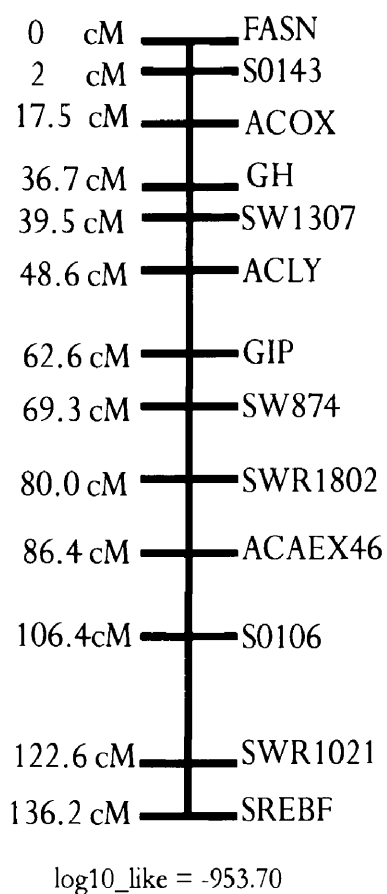
***QTLs y genes candidatos
para composición de ácidos
grasos en grasa subcutánea***

5.1 MAPEO DE LIGAMIENTO

Para construir un nuevo mapa de ligamiento del cromosoma 12 porcino, además de la información empleada en el estudio previo (Muñoz *et al.*, 2007) se incorporó la información del genotipado de los polimorfismos *ACLY*: c.*523T>C y *SREBF1*: c.1023 T>C y de los haplotipos definidos por los polimorfismos *ACOX1*: c.[481G>C; 1238A>G; *21G>A]. En total fueron utilizados siete polimorfismos detectados en genes candidatos (*FASN*, *ACOX1*, *GH*, *ACLY*, *GIP*, *ACACA* y *SREBF1*) y seis microsatélites (*s0143*, *sw1307*, *sw874*, *swr1802*, *s0106* y *swr1021*).

La longitud del nuevo mapa promedio entre ambos sexos fue de 136.2 cM (Figura 5.1) con una distancia media entre los marcadores de 10.48 cM. Los genes *ACOX1*, *ACLY* y *SREBF1* se localizaron en las posiciones 17.5 cM, 48.6 cM y 136.2 cM.

Figura 5.1. Mapa de ligamiento promedio entre ambos sexos del SSC12 construido a partir de los datos de genotipado en los individuos de la F2.



5.2 ANÁLISIS DE DETECCIÓN DE QTLs

Utilizando la información detallada anteriormente se realizó el análisis de detección de QTLs para el porcentaje de los diferentes ácidos grasos en grasa subcutánea así como para los diversos índices calculados (Tabla 3.1). Para ello se utilizó la metodología descrita en el apartado de análisis estadístico de materiales y métodos 3.7.2.

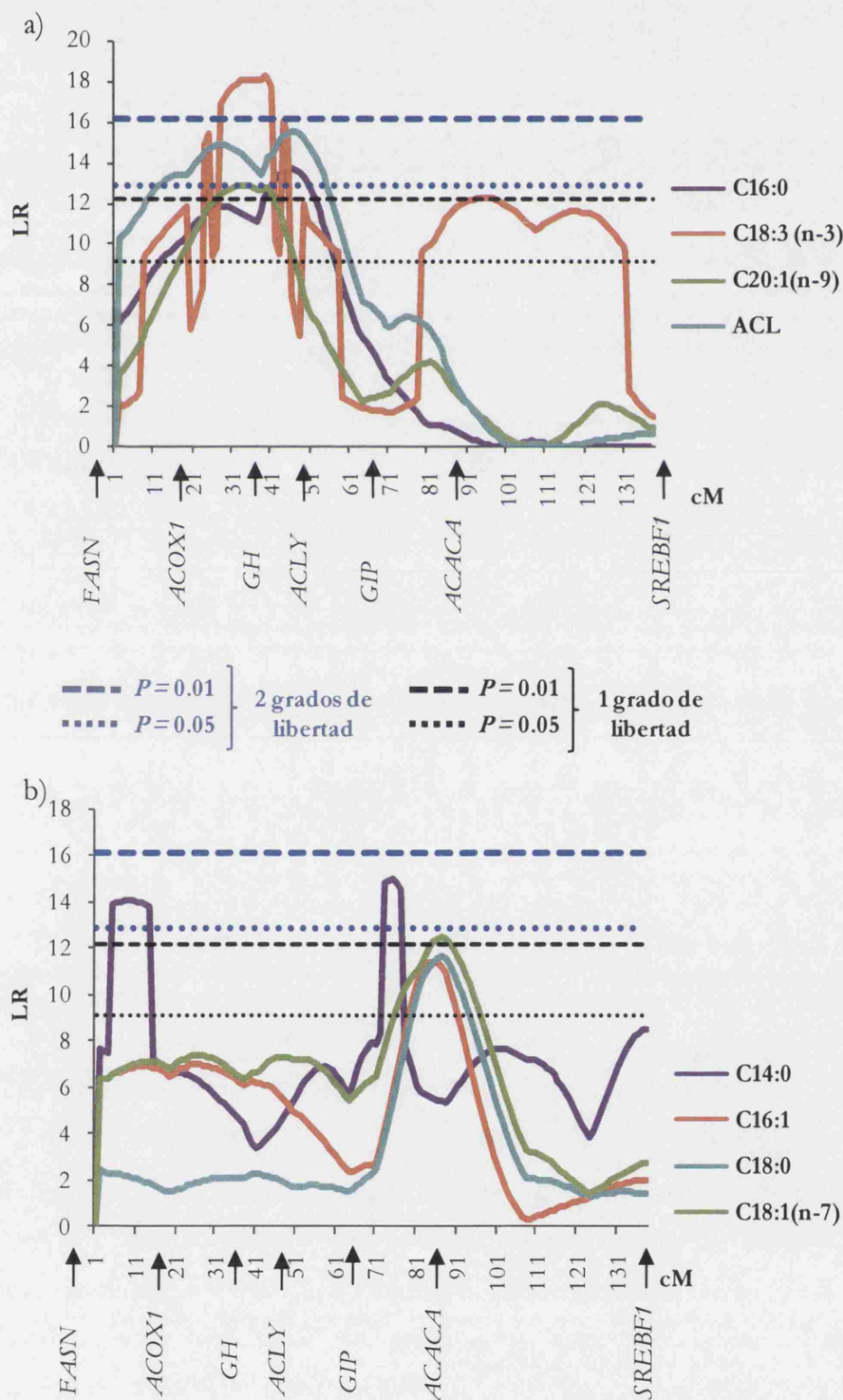
Los umbrales de significación cromosómica calculados, de acuerdo con el número de marcadores y con la longitud total del cromosoma, fueron 19.86, 15.24 y 11.98 en contrastes con dos grados de libertad, asociados con errores de tipo I de 0.1, 1 y 5% respectivamente. Los valores respectivos para contrastes con un grado de libertad fueron 16.49, 12.15 y 9.14.

Un primer examen de los perfiles de la razón de verosimilitud obtenidos en estos análisis reveló dos regiones cromosómicas con presencia de QTLs significativos para algunos de los caracteres citados anteriormente. La primera de estas regiones (QTL1) se localiza en el intervalo situado entre 27 y 52 cM (Figura 5.2 a) y la segunda (QTL2) entre 72 y 92 cM (Figura 5.2 b). Los principales resultados, que se detallan en la tabla 5.1, indican que la mayoría de los efectos detectados son aditivos.

El QTL1 se localiza en la región proximal del cromosoma 12 entre los marcadores *ACOX1* y *ACLY*. El efecto más significativo afecta al porcentaje de ácido linolénico [C18:3 (n-3)] y está situado en la posición 38 cM. Cercano a éste (24 - 41 cM) se ha detectado otro efecto sobre el porcentaje de ácido gadoleico [C20:1(n-9)] así como efectos sobre el porcentaje de ácido palmítico (C16:0) y la Longitud Media de la Cadena (ACL) del total de ácidos grasos, localizados respectivamente en las posiciones de 46 y 45 cM. Los efectos génicos detectados son aditivos para todos los caracteres excepto para el C18:3 (n-3) que presenta efectos aditivos y dominantes significativos. El alelo Q (Ibérico) incrementa el contenido de C16:0 en 1.57 % y produce una disminución en la longitud media de la cadena (ACL), así como en los porcentajes de C20:1 (n-9) y C18:3 (n-3), donde el efecto de la combinación heterocigota (Qq) es más parecido al del genotipo homocigoto (qq) Landrace.

EL QTL2 mapea en la región del cromosoma 12 comprendida entre los marcadores *swr1802* y *ALGA0104951*, en la que se detectaron varios efectos sobre los porcentajes de los ácidos mirístico (C14:0), palmitoleico [(C16:1 (n-9)], esteárico (C18:0) y vaccénico [(C18:1 (n-7)]. El efecto sobre C14:0 es principalmente dominante y el valor máximo de la razón de verosimilitud (LR) está situado en 74 cM. Sin embargo los efectos

Figura 5.2. Perfiles de los diferentes tests de la razón de verosimilitud a lo largo del cromosoma 12. a) Caracteres afectados significativamente por el QTL1. b) por el QTL2.



sobre los otros tres ácidos grasos son aditivos, y los valores máximos de LR están localizados en la posiciones 84 [[C16:1 (n-9)] y 86 cM [C18:0; C18:1 (n-7)], prácticamente coincidentes con la localización del gen *ACACA* (86.4 cM). El alelo *Q* (Ibérico) del QTL2

incrementa el porcentaje de C14:0, C16:1(n-9) y C18:1(n-7) y decrece el contenido del ácido graso C18:0.

Tabla 5.1. Resultados significativos del análisis de detección de QTLs para contenido de ácidos grasos en grasa subcutánea. Posición en el cromosoma 12 e intervalo de confianza del QTL (cM), valor máximo de la razón de verosimilitud (LR), *P*-nominal, efectos aditivo (*a*) y dominante (*d*) y sus errores estándar (entre paréntesis).

Carácter	<i>a</i> (ES)	<i>d</i> (ES)	Posición en cM (IC)	LR	<i>P</i> nominal
C14:0	0.024 (0.013)	-0.050 (0.020)	74 (72-76)	15.02**	5.45 x 10 ⁻⁴
C16:0	0.331 (0.089)	-	46 (39-52)	13.69**	2.16 x 10 ⁻⁴
C16:1 (n-9)	0.096 (0.028)	-	84 (79-89)	11.41*	7.31 x 10 ⁻⁴
C18:0	-0.343 (0.100)	-	86 (80-91)	11.63*	6.48 x 10 ⁻⁴
C18:1 (n-7)	0.090 (0.025)	-	86 (78-92)	12.46**	4.15 x 10 ⁻⁴
C18:3 (n-3)	-0.029 (0.013)	0.052 (0.019)	38 (27-40)	18.33**	1.05 x 10 ⁻⁴
C20:1 (n-9)	-0.060 (0.016)	-	33 (24-41)	12.93**	3.23 x 10 ⁻⁴
ACL	-0.010 (0.002)	-	45 (38-52)	15.54*	8.05 x 10 ⁻⁴

Niveles de significación cromosómica: ** *P* = 0.01; * *P* = 0.05

Además, se realizó un análisis para discernir, en cada una de las citadas regiones cromosómicas, entre la presencia de un solo QTL con efectos pleiotrópicos sobre varios ácidos grasos (hipótesis nula) y la hipótesis alternativa de dos QTLs con efectos sobre diferentes ácidos grasos. Para ello se realizaron contrastes entre modelos de dos QTLs y de un único QTL con efectos pleiotrópicos sobre diversas parejas de ácidos grasos. Estos contrastes sólo se realizaron con los QTLs con efectos exclusivamente aditivos. En todos los casos el valor de *P* nominal obtenido fue inferior al umbral de significación ($\chi^2_{0.05|1} = 3.84$), por lo que no se pudo rechazar la hipótesis nula de un único QTL, en cada una de las regiones del cromosoma, con efectos pleiotrópicos en diferentes ácidos grasos.

A partir de la hipótesis de un modelo de un solo QTL en cada región cromosómica de interés con efectos pleiotrópicos sobre distintos caracteres se testó un modelo multicarácter de efectos aditivos para cada uno de los dos QTLs. Se utilizó un modelo multicarácter con efecto del QTL1 sobre los ácidos grasos C16:0 y C20:1 (n-9) y el índice ACL. Y un modelo multicarácter con efecto del QTL2 sobre C16:1(n-9), C18:0 y C18:1(n-7). El objetivo de estos análisis fue el de precisar la posición de ambos QTLs utilizando

conjuntamente toda la información fenotípica influida por éstos. Los resultados se detallan en la tabla 5.2.

Tabla 5.2. Resultados del análisis multicarácter de detección de QTLs para contenido de ácidos grasos en grasa subcutánea.

QTL	Carácter	Posición en cM	<i>a</i> (ES)	LR	<i>P</i> nominal
	C16:0		0.291 (0.083)		
QTL1	C20:1 (n-9)	40	-0.052 (0.015)	21.81	7.11×10^{-5}
	ACL		-0.010 (0.002)		
	C16:1 (n-9)		0.096 (0.029)		
QTL2	C18:0	86	-0.342 (0.099)	16.49	8.98×10^{-4}
	C18:1 (n-7)		0.091 (0.025)		

5.3 ANÁLISIS DE ASOCIACIÓN DE SNPs EN GENES CANDIDATOS *ACOX1*, *ACLY* y *ACACA*

5.3.1. Modelo animal estándar

En primer lugar, los genes *ACOX1* y *ACLY* fueron evaluados como posibles genes subyacentes al QTL1, localizado entre 27 y 52 cM, que afecta a la longitud media de la cadena y a los ácidos grasos C16:0, C18:3 y C20:1 (Tabla 5.3). En el caso del gen *ACOX1*, se analizó el efecto de los haplotipos H1 (c.[481G; 1238A; *21G]) y H2 (c.[481C; 1238G; *21A]) detectados en el cruce IBMAP. Sólo se utilizó un modelo con aditividad debido a la baja informatividad de estos haplotipos. Sin embargo, para el análisis del polimorfismo *ACLY*: c.*523T>C, con frecuencias intermedias en los animales F2 ($T = 0.43$ y $C = 0.57$), se utilizaron también modelos con aditividad y dominancia. No se detectó ningún efecto significativo de estos polimorfismos o haplotipos sobre los caracteres estudiados, tan sólo efectos aditivos sugestivos del haplotipo H1 del gen *ACOX1* sobre el ácido graso C20:1(n-9), con un valor de 0.045 ± 0.024 ($P < 0.065$), y del SNP *ACLY*: c.*523T>C sobre el ácido graso C16:0, con un valor de 0.182 ± 0.099 ($P < 0.059$). En la siguiente tabla se presentan los resultados obtenidos con el modelo con aditividad ya que los resultados obtenidos con un modelo con aditividad y dominancia eran menos relevantes.

Tabla 5.3. Resultados del análisis de asociación para los haplotipos del gen *ACOX1* y el SNP *ACLY*: c.*523T>C con un modelo animal estándar.

Haplotipo/SNP	Carácter	a (ES)	LR	P
<i>ACOX1</i> : [H1;H2]*	C16:0	-0.137 (0.138)	1.05	0.304
	C18:3	0.016 (0.021)	0.68	0.410
	C20:1(n-9)	0.045 (0.024)	3.40	0.065
	ACL	0.002 (0.004)	0.18	0.670
<i>ACLY</i> : c.*523>C	C16:0	0.182 (0.099)	3.57	0.059
	C18:3	-0.019 (0.015)	1.81	0.180
	C20:1(n-9)	-	0	-
	ACL	-0.004 (0.003)	2.11	0.150

Asimismo, el gen *ACACA* se evaluó como gen candidato para el QTL2, con efectos sobre el porcentaje de C14:0, C16:1(n-9), C18:0 y C18:1(n-7). Los polimorfismos analizados fueron *ACACA*: g.48001 A>G, caracterizado en el presente estudio, y *ACACA*: c.5634T>C, caracterizado previamente por Muñoz *et al.* (2007) (Tabla 5.4). En ambos casos se utilizaron tanto un modelo completo con efectos aditivos y dominantes como otro sólo con efectos aditivos. En el caso del SNP *ACACA*: g.48001A>G se detectó un efecto significativo del polimorfismo sobre el porcentaje del ácido esteárico (C18:0) de 0.248 ± 0.096 ($p < 0.01$). Los resultados más significativos se detectaron en la asociación entre el SNP *ACACA*: c.5634T>C y los porcentajes de C16:1(n-9), C18:0 y C18:1(n-7). El alelo C produce un incremento en los porcentajes de los dos ácidos grasos insaturados y una disminución del porcentaje de C18:0. Estos resultados concuerdan con los obtenidos en el estudio anterior, aunque los efectos estimados en el presente trabajo son de mayor magnitud y más significativos. En ningún caso se encontraron efectos dominantes significativos.

Tabla 5.4. Resultados del análisis de asociación de los SNPs del gen *ACACA*: g.48001 A>G y c.5634T>C con el modelo animal estándar.

SNP	Carácter	a (ES)	LR	P
<i>ACACA</i> : g.48001 A>G	C14:0	-0.084 (0.012)	0.47	0.490
	C16:1(n-9)	-0.038 (0.027)	1.93	0.164
	C18:0	0.248 (0.096)	6.57	0.010
	C18:1(n-7)	-0.033 (0.025)	1.87	0.127
<i>ACACA</i> : c.5634T>C	C14:0	0.009 (0.013)	0.47	0.493
	C16:1(n-9)	0.102 (0.030)	12.23	4.69 x 10 ⁻⁴
	C18:0	-0.331 (0.104)	9.90	0.001
	C18:1(n-7)	0.093 (0.026)	13.05	3.03 x 10 ⁻⁴

5.3.2. Modelo animal de asociación asistida por marcadores

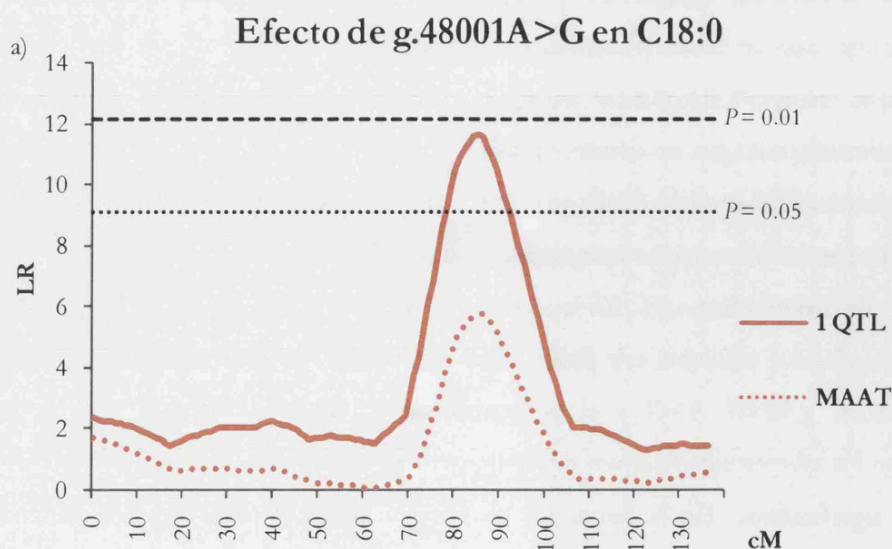
Es sabido ya que los cruces F2 entre líneas divergentes facilitan la detección de QTL (Varona *et al.*, 2005) pero dificultan la distinción entre mutaciones causales y neutras dado el fuerte desequilibrio de ligamiento generado en dichos cruces. El test de asociación asistida por marcadores (apartado 3.7.3.2) se realizó con el objetivo de reducir la proporción esperada de falsos positivos. El análisis de asociación con este modelo más complejo, que incorpora los efectos del QTL y del polimorfismo, se llevó a cabo para aquellos polimorfismos que resultaron significativos en el análisis de asociación anterior. Para ello se contrasta el modelo dónde se contemplan los efectos del SNP y del QTL frente a cada uno de los modelos que contemplan sólo uno de los efectos.

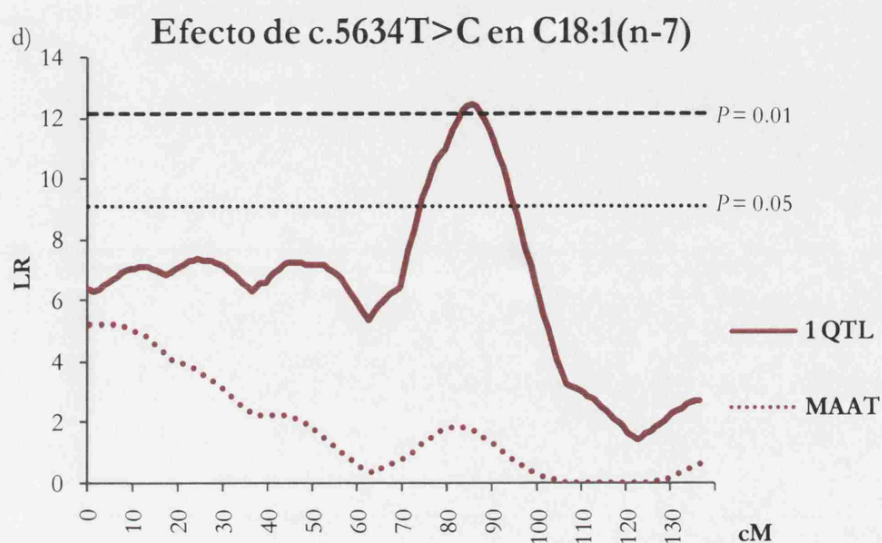
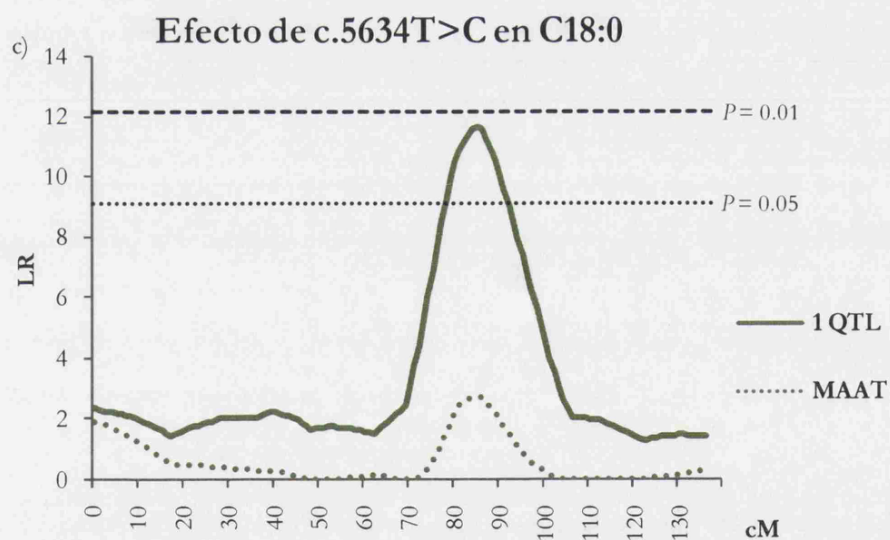
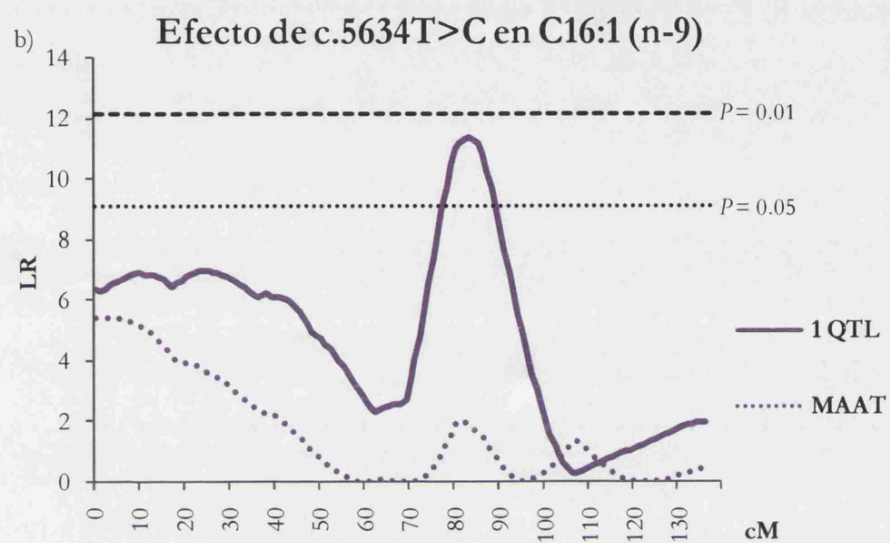
Los polimorfismos *ACACA*: g.48001A>G y *ACACA*: c.5634T>C fueron igualmente analizados mediante este procedimiento (Tabla 5.5). Como resultado del análisis entre *ACACA*: g.48001 A>G y el porcentaje de C18:0, en el primer contraste se mantuvieron los efectos significativos del SNP, pero en el segundo, los efectos del SNP ya no fueron significativos. En la figura 5.3 se observa como el perfil del LR disminuye cuando se incluye el efecto del SNP. Cuando se analizó el polimorfismo *ACACA*: c.5634T>C con este modelo se obtuvieron resultados más concluyentes que con el anterior polimorfismo. El efecto del SNP es altamente significativo, y además se reducen o desaparecen los efectos del QTL correspondiente en su primitiva posición (Figura 5.3).

Tabla 5.5. Contraste del efecto de los SNPs en el gen *ACACA*, g.48001A>G y c.5634T>C analizados en el test de asociación por marcadores.

Efecto del SNP					QTL			
SNP	Carácter	LR	P	a (ES)	cM	LR	P nominal	a (ES)
g.48001A>G	C18:0	5.76	0.0164	0.237 (0.098)	86	5.73	0.0166	-0.260 (0.108)
	C16:1(n-9)	13.27	2.70×10^{-4}	0.120 (0.032)	3	5.44	0.0197	0.062 (0.027)
c.5634T>C	C18:0	9.04	2.64×10^{-4}	-0.317 (0.105)	85	2.70	0.099	-0.242 (0.147)
	C18:1(n-7)	11.53	0.0068	0.092 (0.027)	1	5.22	0.022	0.053 (0.023)

Figura 5.3. Perfil de la razón de verosimilitud de QTL cuando se incorpora en el modelo el polimorfismo correspondiente. Las líneas continuas representan el análisis con el modelo de un QTL y las discontinuas el análisis con el modelo que incluye además el SNP correspondiente. Perfiles correspondientes a los análisis: a) g.48001A>G/C18:0 b) c.5634T>C/C16:1(n-9) c) c.5634T>C/C18:0 d) c.5634T>C/C18:1(n-7).





Este resultado indica que el desequilibrio con la mutación causal es mayor, en el caso del SNP, que en el del QTL. Para contrastar la causalidad del SNP es necesario comprobar la consistencia de sus efectos en otras poblaciones y realizar estudios de expresión y/o funcionales.

5.4. OTROS ANÁLISIS DE ASOCIACIÓN

5.4.1. Modelo de tres alelos.

Este modelo asume que el polimorfismo, en este caso el SNP *ACACA* c.5634T>C, está ligado a la mutación causal solamente en una de las poblaciones parentales. En el modelo se distinguen tres alelos: un alelo C asociado al origen Ibérico (C_I), un segundo alelo C asociado al origen Landrace (C_{LR}) y un tercero alelo T de origen exclusivo Landrace. El modelo general se ha descrito en Material y Métodos (3.6.3). El modelo con tres alelos se ha denominado *M5a* y el de dos alelos *M3a*.

En la tabla 5.6 se describen los resultados del análisis de asociación con el modelo con tres alelos. Donde a_I es el efecto de sustitución del alelo Ibérico y a_{LR} el efecto de sustitución del alelo Landrace. El análisis de asociación sólo fue realizado para aquellos ácidos grasos para los que se detectó QTL con efecto sobre los mismos en una posición cercana a la localización del gen *ACACA*. En la última columna se comparan los modelos *M3a* vs *M5a*.

La comparación entre estos modelos mediante la razón de verosimilitud indica que los resultados no aportan evidencia de que el modelo de tres alelos explique mejor los resultados que el test de asociación mediante el modelo animal estándar con dos alelos.

Tabla 5.6. Resultados de los análisis de asociación realizados para el polimorfismo *ACACA* c.5634T>C con un modelo de tres alelos.

Caracter	LR	P	a_I (ES)	a_{LR} (ES)	<i>M5a</i> vs <i>M3a</i>
C14:0	0.87	0.647	0.012 (0.014)	0.002 (0.002)	0.40
C16:1	13.64	0.005	0.112 (0.030)	0.062 (0.040)	1.41
C18:0	12.18	0.002	-0.373 (0.107)	-0.178 (0.151)	2.28
C18:1 (n-7)	14.34	7.68×10^{-4}	0.103 (0.027)	0.057 (0.039)	1.29

$\chi^2_{0.05|11} = 3.84$ y $\chi^2_{0.01|11} = 6.64$.

5.4.2. Test de asociación con haplotipos.

Los genotipos de los polimorfismos del gen *ACACA* (g.48001A>G y c.5634T>C) analizados en este trabajo permiten definir cuatro haplotipos en el cruce IBMAP: H1 [g.48001A; c.5634C], H2 [g.48001G; c.5634C], H3 [g.48001A; c.5634T], H4 [g.48001G; c.5634T]. Las frecuencias haplotípicas de estas combinaciones en los animales de la F2 fueron respectivamente: 0.35, 0.27, 0.34 y 0.04.

En la tabla 5.7 se muestran los resultados del test de asociación que al igual que en el análisis anterior sólo se aplicó para aquellos ácidos grasos para los que se detectó QTL con efecto sobre los mismos en una posición cercana a la localización del gen *ACACA*. En este test se compara un modelo de haplotipos (*M6a*) frente a un modelo de dos alelos (*M2a*). En la última columna de la tabla 5.7 se refleja esta comparación que no supera los umbrales de distribución χ^2 con dos grados de libertad (5.99 y 9.21 al 5 y 1% respectivamente). Por lo tanto este modelo no explica mejor los resultados que el test de asociación con dos alelos.

Tabla 5.7. Resultados de los análisis de asociación realizados con las combinaciones haplotípicas de los polimorfismos g.48001A>G y c.5634T>C.

Caracter	LR	P	<i>M6a</i> vs <i>M2a</i>
C14:0	2.33	0.507	0.73
C16:1	11.73	0.008	-
C18:0	14.18	0.003	2.14
C18:1(n-7)	13.34	0.004	0.15
$\chi^2_{05[2]} = 5.99$ y $\chi^2_{01[2]} = 9.21$			

5.5. ANÁLISIS DEL POLIMORFISMO *ACACA* c.5634T>C EN CERDOS IBÉRICOS PUROS Y CRUZADOS CON DUROC

Los efectos del polimorfismo *ACACA*: c.5634T>C sobre el perfil de ácidos grasos de grasa subcutánea recogidos en las tablas 5.4 y 5.5 fueron evaluados en animales pertenecientes a un núcleo de selección de cerdos Ibéricos, así como en animales

procedentes de cruces de cerdas del citado núcleo con verracos Duroc. Otros detalles de este ensayo se han descrito en Material y Métodos.

5.5.1. Animales *Ibérico x Ibérico*

Un total de 174 animales Ibéricos puros fueron genotipados para el polimorfismo *ACACA*: c.5634T>C. El alelo más frecuente fue el alelo T (0.68) a diferencia de lo observado en los parentales ibéricos del cruce ILMAP, en los que aparece fijado el alelo alternativo C.

Como resultado del análisis de asociación entre el SNP *ACACA*: c.5634T>C y el perfil de ácidos grasos de grasa subcutánea mostrado en la tabla 5.8 se detectó asociación significativa entre el polimorfismo y los ácidos grasos C12:0 y C14:0. La sustitución del alelo T por el C produce un incremento del porcentaje de ambos ácidos grasos. En el cruce ILMAP no se detectaron efectos de este SNP sobre estos ácidos grasos. Por el contrario no se detectó asociación significativa sobre los ácidos grasos C16:1(n-7), C18:0 y C18:1(n-7), para los cuales, en el cruce ILMAP, el alelo *ACACA* c.5634C producía un efecto negativo sobre el ácido graso C18:0 y positivo sobre los dos restantes.

Tabla 5.8. Resultados más relevantes del análisis de asociación entre el polimorfismo *ACACA*: c.5634T>C y el perfil de ácidos grasos en grasa subcutánea.

Carácter	LR	P	a_{SNP} (ES)
C12:0	5.50	0.019	0.002 (0.001)
C14:0	5.42	0.020	0.036 (0.015)
C16:1 (n-9)	1.16	0.281	0.042 (0.039)
C18:0	0.04	0.844	-0.026 (0.132)
C18:1(n-7)	1.31	0.253	-0.211 (0.184)

5.5.2. Animales *Duroc x Ibérico*

Un total de 107 animales cruzados fueron genotipados para el SNP *ACACA*: c.5634T>C. Se observaron frecuencias alélicas intermedias en esta población (T = 0.48; C = 0.52). En el análisis de asociación no se encontró ningún efecto significativo del polimorfismo sobre los ácidos grasos analizados.

5.6. ANÁLISIS DE EXPRESIÓN DEL GEN *ACACA*

La siguiente tarea realizada para validar los efectos de los SNPs del gen *ACACA* sobre la composición de ácidos grasos de grasa subcutánea fue analizar la expresión de este gen en muestras de este mismo tejido. Lo ideal hubiese sido realizar el análisis de expresión en muestras de los mismos animales pertenecientes a la generación F2 en donde se hizo la toma de datos fenotípicos y los análisis anteriormente expuestos. Sin embargo no se dispone de muestras de tejidos de los animales F2 para extracción de ARN y el estudio se hizo en la generación RC2 del ILMAP, que tiene los mismos padres (machos F1) que la F2 y muestras de tejido disponibles.

Para cuantificar la expresión relativa del gen se empleó la técnica de PCR cuantitativa a tiempo real con “SYBR green” y se utilizaron dos genes control.

5.6.1. Selección de genes control

Los genes control se expresan de manera constante bajo cualquier condición universal en una muestra problema. Además, es conveniente que el gen endógeno presente niveles de expresión y eficiencia de amplificación similares a los del gen problema que se quiere cuantificar. Hasta la fecha no se conoce ningún gen endógeno que sea universal y válido para todos los tejidos en distintas condiciones experimentales. Por ello, en cada experimento hay que valorar si los genes seleccionados son adecuados y verificar que tengan la menor variación posible en la expresión.

Tabla 5.9. Parejas de oligonucleótidos empleados para cuantificar los genes control, tamaño del fragmento amplificado (pb) y temperatura de hibridación (Ta)

Gen	Secuencia del oligonucleótido	Pb	Ta
<i>B2MF</i>	TTCACACCGCTCCAGTAG	166	60
<i>B2MR</i>	CCAGATACATAGCAGTTCAGG		
<i>GAPDHF</i>	TCGGAGTGAACGGATTTG	219	60
<i>GAPDHR</i>	CCTGGAAGATGGTGATGG		

Basado en una búsqueda bibliográfica se seleccionaron dos genes para usarlos como control, el gen *GAPDH* y el gen *B2M*. Las secuencias de las parejas de oligonucleótidos obtenidas de Kuijk *et al.* (2007) se muestran en la tabla 5.9.

Se realizó una mezcla del ADNc de todos los individuos a ser valorados y, a partir de esta, se prepararon diluciones seriadas 1/20, 1/40, 1/80 y 1/160. Se utilizaron las cuatro diluciones por triplicado, para disminuir el error de pipeteo, se midió la expresión de los genes control y se realizó la curva estándar para medir la eficiencia de la PCR.

Calculados los Ct y las cantidades relativas de cada uno de los genes control, se utilizó la herramienta informática geNorm (Vandesompele *et al.*, 2002) para testar si los genes seleccionados eran adecuados como control. Se ha utilizado el índice de estabilidad (M), aceptándose como óptimos valores por debajo de 1.5. (Tabla 5.10.) y, como ambos genes tuvieron un M inferior a este valor se aceptaron como genes control. Además se obtuvieron valores de la eficiencia de la PCR comprendidos entre -3.1 y -3.6, el rango considerado óptimo para normalizar los datos de expresión.

Tabla 5.10. Resultado de la eficiencia de la PCR de los genes control.

Gen	Eficacia (E)	Estabilidad (M)
<i>B2M</i>	-3.308	0.545
<i>GAPDH</i>	-3.277	0.545

5.6.2. Estudio de la expresión del gen *ACACA*

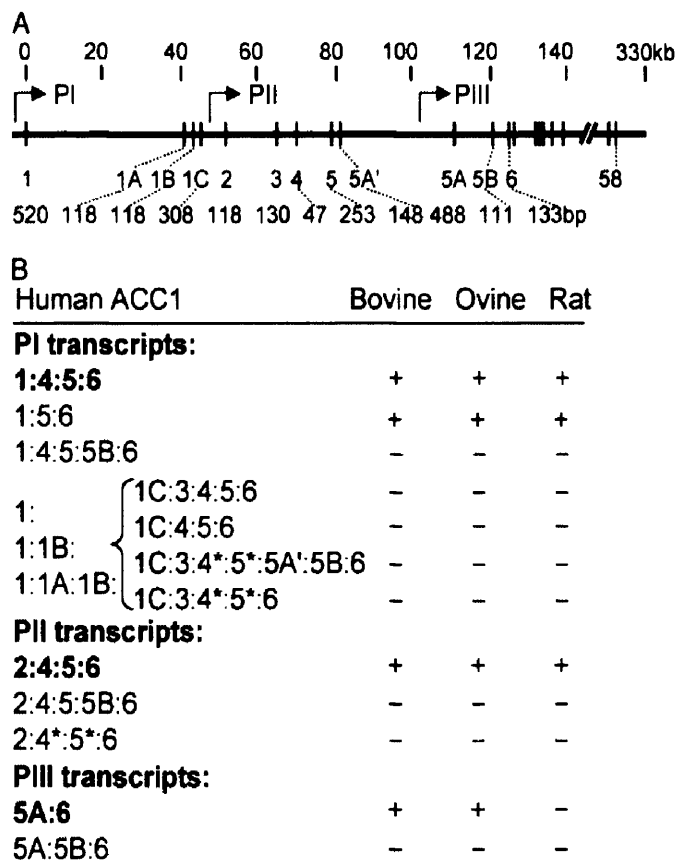
Los animales fueron clasificados en función de su genotipo para los polimorfismos del gen *ACACA* g.48001A>G y c.5634T>C. Se eligieron treinta animales de manera que hubiese diez de cada genotipo de ambos polimorfismos y se valoró la expresión del gen *ACACA* en muestras de grasa dorsal.

5.6.2.1. Diseño de oligonucleótidos para PCR cuantitativa a tiempo real.

Según se ha descrito en varias especies la expresión del gen *ACACA* es bastante compleja (Mao *et al.*, 2003). Los transcritos expresados difieren fundamentalmente en la secuencia de la zona 5'UTR (Figura 5.4). Por lo tanto, con el objetivo de detectar la máxima cantidad de expresión se diseñó una pareja de oligonucleótidos localizada en una zona común a la mayoría de transcritos que se pudiesen detectar en la grasa dorsal. Esta pareja

estaba localizada entre los exones 5 y 6 y presentó un nivel óptimo con un Ct alrededor de 27. Los oligonucleótidos diseñados, ACARTF (5'-CTGAGAGCTCGTTTTGAAGGAATA-3') y ACARTR (5'-TTTACTAGGTGCAAGCCAGACAT-3'), permitieron amplificar un fragmento de 281 pb con un valor elevado de eficiencia de PCR (-3.226).

Figura 5.4. Figura extraída de Mao *et al.* (2003) donde se representa la heterogeneidad de la región 5'-ARNm del gen *ACACA*. A) Organización de la región 5'UTR humana. Cada barra vertical representa un exón y los promotores se designan por PI, PII y PIII. B) Estructura de los transcritos obtenidos a partir de cada uno de los tres promotores y comparación con sus homólogos en bovino, ovino y rata.



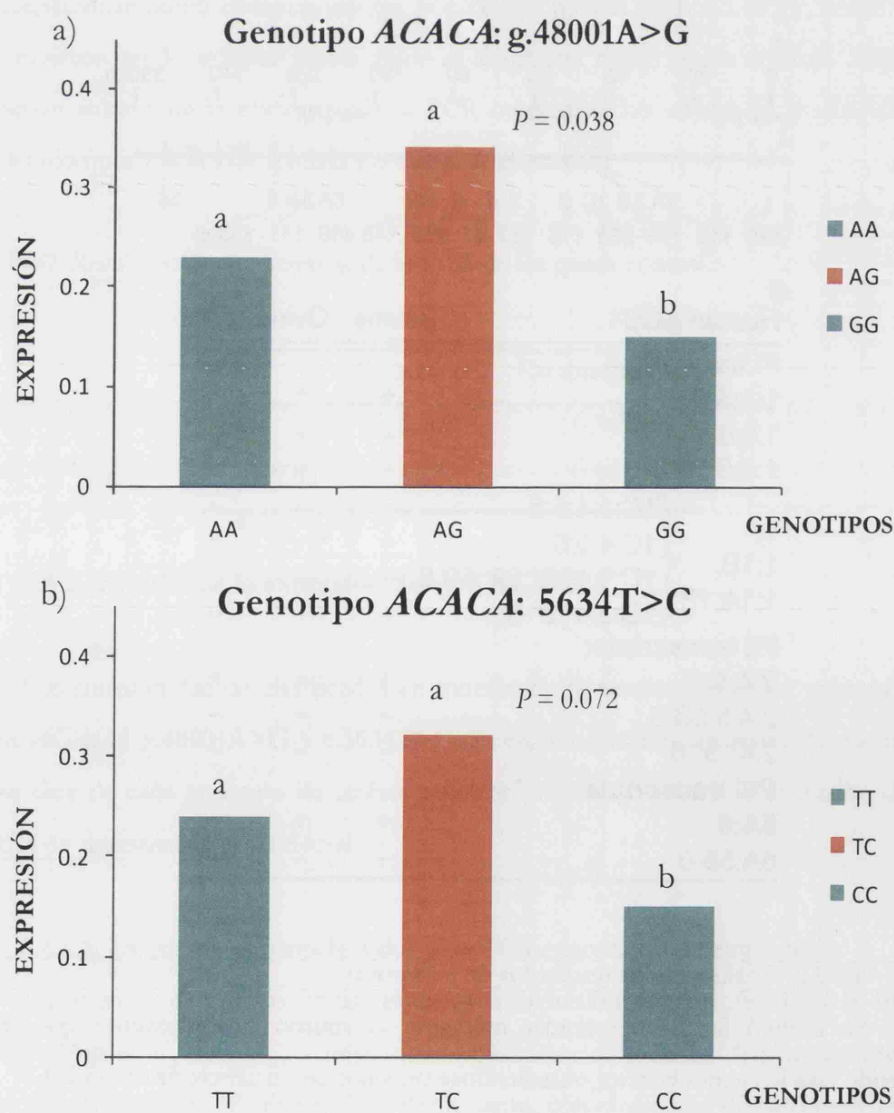
5.6.2.2. Análisis de los resultados de expresión

Se calculó la curva estándar mediante el mismo procedimiento que se había empleado con los genes control obteniéndose un valor de eficiencia de $m = -3.226$.

Como resultado del análisis de expresión del gen *ACACA* clasificado en base a los genotipos del polimorfismo g.48001A>G se detectó un efecto significativo de estos genotipos ($P < 0.038$). La comparación múltiple mediante el test de Tukey, mostró que los

individuos homocigotos AA y heterocigotos AG presentan una mayor expresión del transcrito que los homocigotos GG (Figura 5.5.a). En el caso de la clasificación en función de los genotipos del polimorfismo *ACACA*: c.5634T>C, se detectó un efecto sugestivo ($P < 0.072$). Como se puede comprobar en la figura 5.5.b, la expresión de los individuos homocigotos CC es menor que la de los individuos con otros genotipos.

Figura 5.5. Expresión en grasa dorsal del gen *ACACA*. a) Clasificación en función del genotipo g.48001A>G. b) Clasificación en función del genotipo c.5634T>C. Clases con letras distintas (a y b) corresponden a valores de expresión significativamente diferentes.



***QTLs y genes candidatos
para composición de ácidos
grasos en grasa intramuscular***

6.1. ANÁLISIS ESTADÍSTICOS EN GRASA INTRAMUSCULAR

El contenido en grasa intramuscular y su composición de ácidos grasos fueron medidos en muestras de *longissimus dorsi* procedentes de animales de las generaciones F3, RC1 y RC2. El total de la información fenotípica disponible se ha resumido en la tabla 3.2. Todos estos individuos con datos fenotípicos, así como sus progenitores en las generaciones anteriores (F0, F1 y F2), fueron genotipados con el *PorcineSNP60 Genotyping BeadChip*, así como para los microsatélites y SNPs de genes candidatos utilizados en el análisis previo de la composición de grasa subcutánea.

Del total de 62163 sondas incluidas en el citado *PorcineSNP60*, sólo 1.197 están localizadas en el SSC12 (build 9). Los SNPs útiles para el análisis del SSC12 fueron 978, una vez descontados aquellos en que no fue posible discriminar de modo unívoco los tres genotipos, además de aquellos poco informativos por presentar una frecuencia del alelo menos frecuente (MAF) inferior a 0.15.

A partir de estos datos se realizaron varios análisis empleando diferente número y tipo de marcadores genéticos. En el primero de ellos se combinó información de 26 SNPs elegidos entre los genotipados con el *chip SNP60* con la información de seis microsatélites (análisis con información combinada) para obtener resultados fácilmente comparables con los obtenidos en el análisis de grasa subcutánea. En el posterior análisis se utilizó únicamente la información proporcionada por el *chip SNP60*.

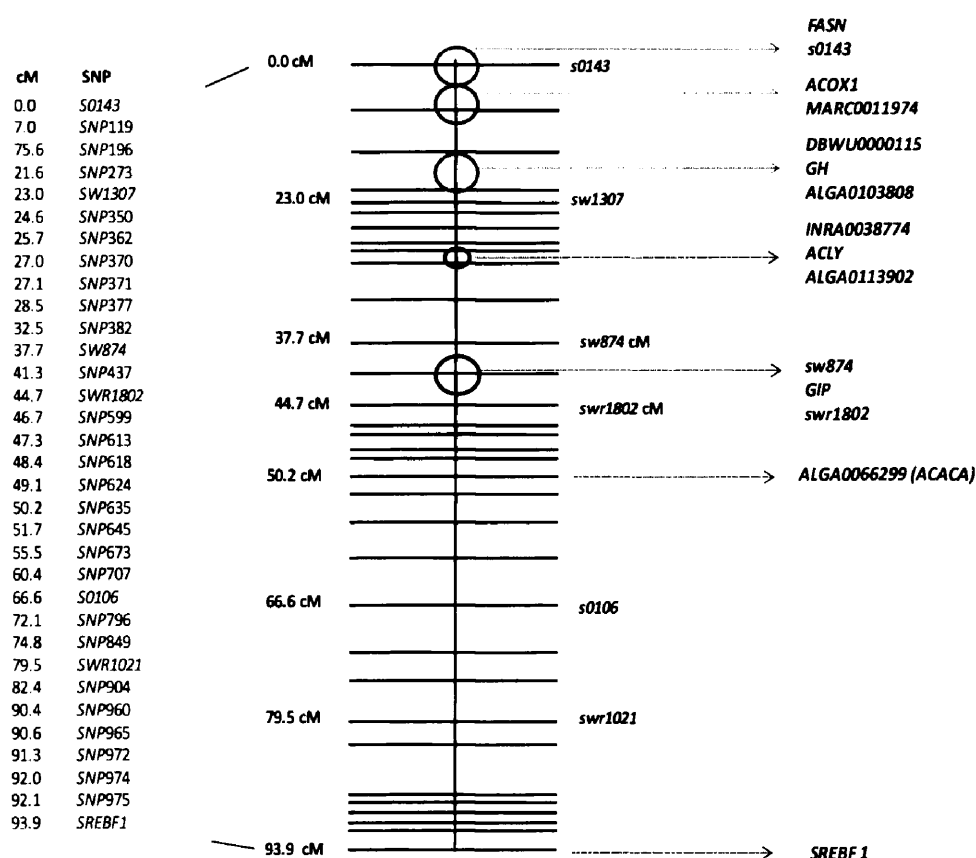
6.2. NUEVOS MAPAS DE LIGAMIENTO DEL CROMOSOMA 12

6.2.1. Análisis con información combinada

Para construir un segundo mapa de ligamiento del cromosoma 12, se utilizaron los datos de genotipado en animales pertenecientes a las generaciones F3, RC1 y RC2 así como en sus progenitores en las generaciones anteriores (F0, F1 y F2) empleando seis microsatélites (*s0143*, *sw1307*, *sw874*, *swr1802*, *s0106* y *swr1021*), el SNP del gen candidato *SREBF1*: c.1023 T>C y 26 SNPs del *chip SNP60*, distribuidos a lo largo del cromosoma y con MAF superior a 0.30. En total fueron empleados 33 marcadores genéticos para la construcción de este mapa. En una primera aproximación, también fueron incluidos SNPs de otros genes candidatos funcionales (*FASN*, *ACOX1*, *GH*, *ACLY*, *GIP*, *ACACA*), pero fueron excluidos ya que no fue posible separar su posición de la de los marcadores

adyacentes. La longitud del nuevo mapa promedio entre ambos sexos fue de 93.9 cM (Figura 6.1) con una distancia media entre los marcadores de 3.02 cM.

Figura 6.1. Mapa genético promedio entre ambos sexos del SSC12 construido a partir de los datos de genotipado en los individuos de la F3, RC1 y RC2 y de sus progenitores en las generaciones anteriores (F0, F1 y F2). Las flechas de colores, en el lado derecho de la figura señalan la ubicación teórica de los genes candidatos.



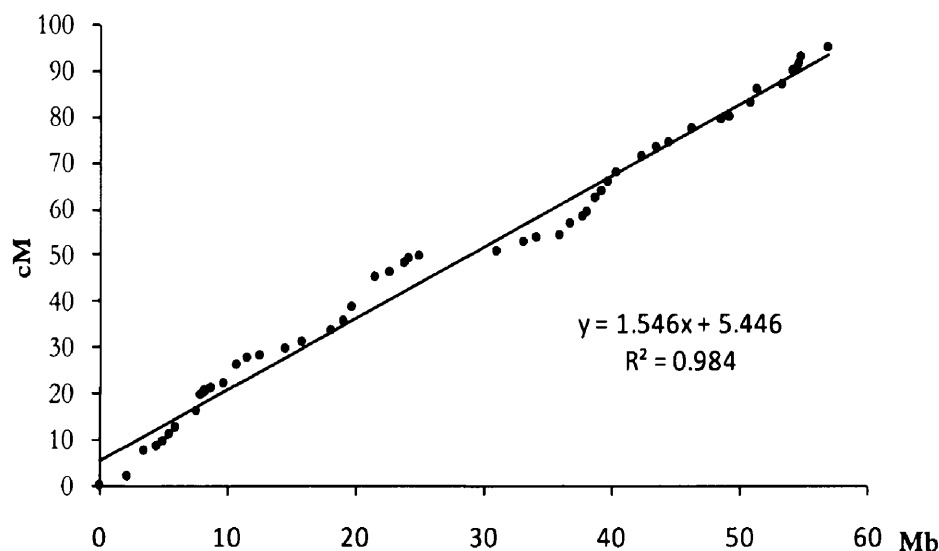
En las bases de datos Ensembl (build9) y NCBI están disponibles las posiciones físicas de los genes candidatos: *ACOX1*, *GH*, *ACLY* y *ACACA*, y por tanto la posición teórica de éstos en el mapa fue inferida a partir de su localización física y de la de los SNPs del *PorcineSNP60 Genotyping BeadChip* empleados en la realización del mapa de ligamiento. Como no se disponía de la localización física del gen *GIP* se ha deducido su posición teórica a partir del mapa comparativo humano-cerdo (www.animalgenome.org/cgi-bin/QTLdb/SS/back/draw_rh2human).

6.2.2 Análisis con 50 SNPs: equivalencia entre mapas físico y genético del SSC12

Del total de los 978 SNPs útiles, y por limitación del software disponible para la detección de QTL, se seleccionaron uno de cada dos sucesivos para ser incluidos en el análisis basado exclusivamente en genotipos obtenidos con el *chip* SNP60. Por lo tanto, en este nuevo análisis se utilizó un total de 489 SNPs.

Debido a la imposibilidad de realizar un mapa genético con 489 marcadores en el momento de elaborar esta tarea, se seleccionaron 50 SNPs de ellos. Como en el caso anterior, el objetivo de la realización de este mapa fue relacionar las distancias física y de ligamiento entre los marcadores seleccionados. La complejidad del mapa obtenido no se muestra en detalle. La longitud del mapa promedio obtenido fue de 95.2 cM y la distancia media entre marcadores de 1.90 cM. Con las posiciones físicas y genéticas de estos 50 marcadores se calculó la regresión lineal entre ambas, representada en la figura 6.2.

Figura 6.2. Recta de regresión de la posición genética (cM) sobre la posición física (Mb) de los 50 marcadores seleccionados.



Según esta recta de regresión, basada en 50 SNPs mapeados genéticamente, 1Mb sería equivalente a 1.546 cM. Por lo tanto, si en el SSC12 se dispone de 978 marcadores en 56.92 Mb, en 1 Mb (1.546 cM) hay un promedio de 17.18 SNPs y en 1 cM habrá 11.11

marcadores. Aunque la regresión lineal presenta un buen ajuste ($R^2=0.984$), en la región central del cromosoma se observa una clara desviación de la regresión estimada.

6.3. ANÁLISIS DE DETECCIÓN DE QTLs.

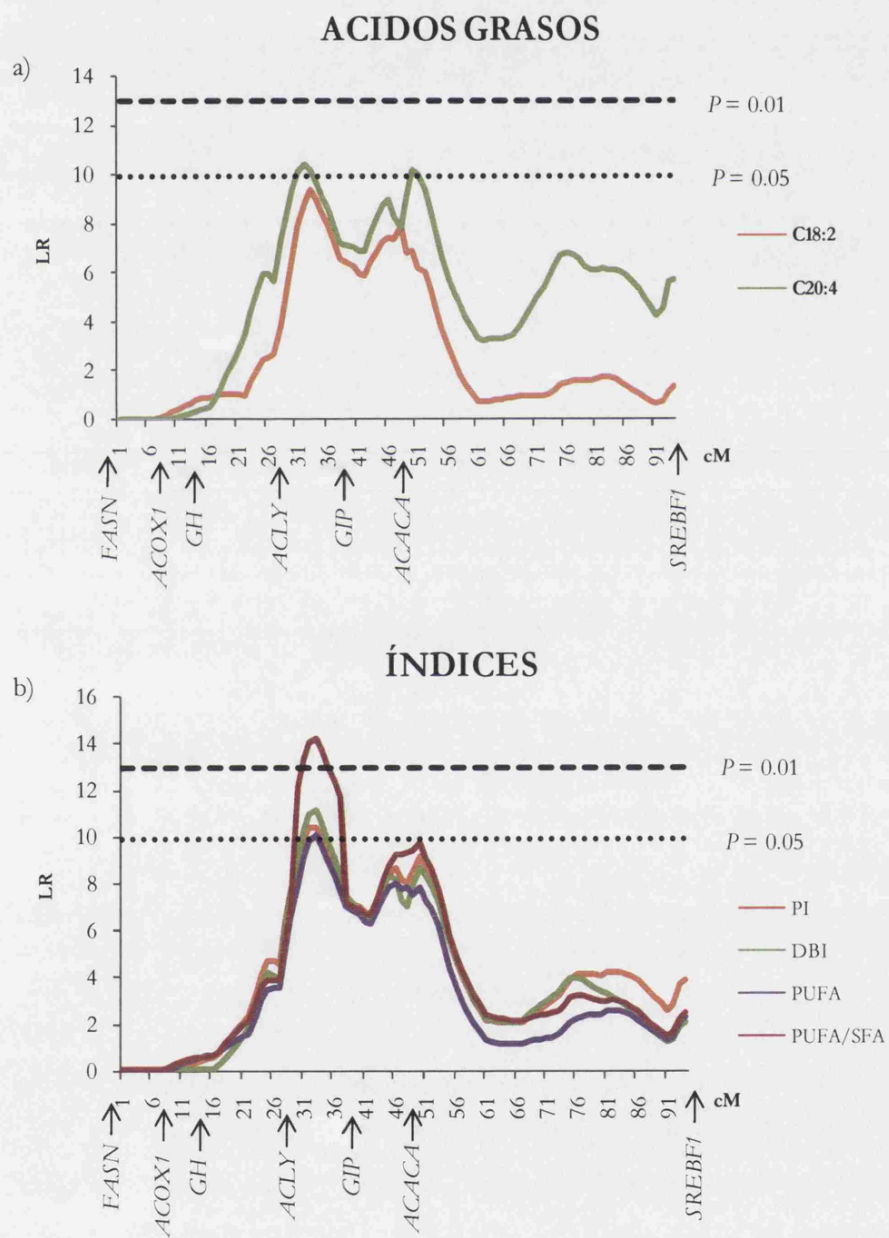
6.3.1 Análisis con información combinada

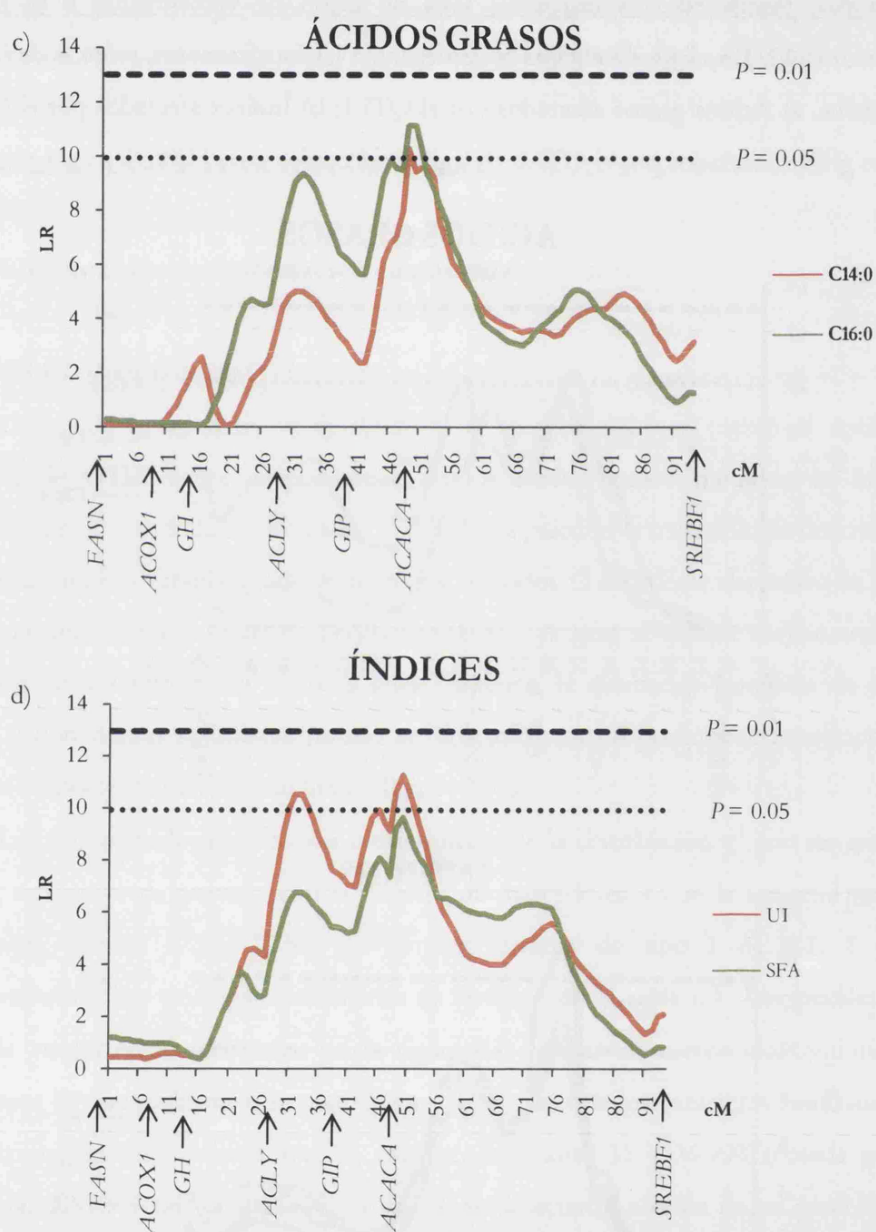
6.3.1.1 Modelo de un QTL

Utilizando la información detallada en el epígrafe 6.2.1 se realizó el análisis de detección de QTLs sobre el porcentaje de los ácidos grasos medidos en la grasa intramuscular y de los índices calculados (Tabla 3.2) siguiendo la metodología descrita en el apartado de análisis estadístico de materiales y métodos (3.6.2). Esta detección se realizó calculando únicamente los efectos genéticos aditivos ya que, al utilizar mayoritariamente individuos de dos retrocruces sobre la línea Landrace, la estimación conjunta de efectos aditivos y dominantes se hubiera basado en una información claramente insuficiente del genotipo homocigoto de alelos de origen Ibérico (QQ).

Los umbrales de significación cromosómica de la distribución χ^2 con un grado de libertad, calculados de acuerdo con el número de marcadores y con la longitud total del cromosoma, fueron 17.34, 12.98 y 9.95 para errores de tipo I de 0.1, 1 y 5% respectivamente. Los principales resultados se resumen en la tabla 6.1. Los perfiles de la razón de verosimilitud mostrados en la figura 6.3 revelan al menos dos regiones del cromosoma 12 que pudieran sustentar QTLs para algunos de los caracteres analizados. En la región proximal del cromosoma 12, comprendida entre 31 y 36 cM acotada por los marcadores *SNP377* (*ALGA0113902*) y *sw874*, se detectaron efectos de un posible QTL (QTL1) sobre varios ácidos grasos e índices. La posición inferida de los genes *GIP* y *ACLY* les sitúa próximos a este intervalo. El mayor efecto de este QTL corresponde al cociente PUFA/SFA y se localiza en la posición 33 cM. En esta misma posición se han detectado efectos significativos sobre los índices PUFA, DBI, PI y sobre el ácido graso C18:2, aunque éste último con un nivel de significación cromosómica próximo al 5% (LR = 9.42). Asimismo, en la posición 32 cM se localiza un QTL con efecto sobre el porcentaje del ácido graso C20:4. En la mayoría de estos casos, el alelo ibérico *Q* tiene un efecto negativo, excepto en el caso del ácido graso C18:2, en el que produce un incremento en el porcentaje del mismo (0.090 ± 0.025).

Figura 6.3. Perfiles de los diferentes tests de razón de verosimilitud a lo largo del cromosoma 12: En el eje de abcisas se representan las localizaciones teóricas de los genes candidatos. a) Ácidos grasos afectados por el QTL1; b) Índices afectados por el QTL1; c) Ácidos grasos afectados por el QTL2; d) Índices afectados por el QTL2.





Por otra parte, en la región del cromosoma 12 comprendida entre 45 y 53 cM, acotada por los microsatélites *sw874* y *s0106*, se detectaron efectos aditivos (QTL2) sobre los porcentajes de los ácidos grasos C14:0 y C16:0 y sobre los índices UI y SFA. En el caso del C14:0, el valor máximo de la razón de verosimilitud está situado en 49 cM mientras los demás están en la posición inmediata (50 cM). Dentro de este intervalo está situado el gen *ACACA*. El alelo *Q* del QTL2 produce un incremento en la cantidad de ácidos grasos saturados, reflejado en los porcentajes de C14:0 y C16:0 y SFA y un efecto negativo sobre el índice UI.

Tabla 6.1. Resultados más significativos del análisis de detección de QTL en grasa intramuscular. Localización en cM e intervalos de confianza (IC), valor de la razón de verosimilitud (LR), valor nominal de P y efecto aditivo (a) y su error estándar (ES) entre paréntesis.

Carácter	Posición en cM (IC)	a (ES)	LR	P -nominal
C14:0	49 (49-53)	0.073 (0.023)	10.30*	0.0013
C16:0	50 (45-52)	0.633 (0.188)	11.12*	8.53×10^{-4}
C18:2	33 (31-36)	0.090 (0.025)	9.42*	0.0021
C20:4	32 (31-34)	-0.461 (0.142)	10.39*	0.0012
PI	33 (31-34)	-2.770 (0.855)	10.39*	0.0013
DBI	33 (31-35)	-0.033 (0.010)	11.15*	8.41×10^{-4}
UI	50 (45-52)	-0.123 (0.037)	11.26*	7.93×10^{-4}
SFA	50 (49-53)	0.777 (0.249)	9.58*	0.0020
PUFA	33 (31-36)	-1.444 (0.453)	10.08*	0.0015
PUFA/SFA	33 (31-36)	-0.047 (0.014)	14.19**	1.65×10^{-4}

** $P = 0.01$; * $P = 0.05$

Los perfiles de la razón de verosimilitud en el QTL1 tienen una forma similar para la mayoría de los caracteres analizados con localizaciones coincidentes del máximo valor de LR o que no distan más de 1 cM. Resultados similares también se observan para el QTL2. En función de lo expuesto, parece probable que haya en cada región un QTL con efecto sobre varios caracteres. Para verificar esta hipótesis se testó para el QTL1 un modelo multicarácter, con efectos sobre los ácidos grasos C20:4 y C18:2 y en el caso del QTL2 un modelo multicarácter de un QTL con efectos sobre los ácidos grasos C14:0 y C16:0. No se han incluido los índices en ninguno de los modelos para evitar el uso de información redundante ya que éstos fueron calculados a partir de los porcentajes de los ácidos grasos. En la tabla 6.2 están expuestos los resultados obtenidos de este análisis. Estos valores no superaron el umbral de significación ($\chi^2_{0.05|11} = 3.84$) en ninguno de los casos. Para C18:2 y C20:4 la comparación de las razones de verosimilitud toma un valor de 0.92 (10.34 – 9.42) y para C14:0 y C16:0, un valor de 0.51 (11.63 – 11.12). Por lo tanto, a pesar de lo esperado, el modelo multicarácter no explica mejor los resultados, pudiendo deberse a que en ambos casos el efecto sobre uno de los dos ácidos grasos es notablemente inferior que el otro.

Tabla 6.2. Resultados del análisis multicarácter de un QTL para cada una de las regiones.

	Carácter	Posición en cM (IC)	a (ES)	LR	P -nominal
QTL1	C18:2	33	-0.947 (0.025)	10.34	0.006
	C20:4		-0.461 (0.142)		
QTL2	C14:0	49	-0.196 (0.065)	11.63	0.003
	C16:0		-1.168 (0.538)		

6.3.1.2 Modelo de dos QTLs.

Los perfiles de LR (figura 6.3), indican que algunos de los caracteres afectados por el QTL1 presentan un segundo máximo local en la zona del QTL2 mientras otros de los afectados por el QTL2, tienen un segundo máximo local en la zona del QTL1. En función de estos resultados, parece justificada la comparación entre los modelos de dos QTLs y de un único QTL. Para ello se testó un modelo con dos QTLs ($M2a$) frente al modelo nulo ($M0$) y frente a un modelo de 1 QTL ($M1a$). Los resultados mostrados en la tabla 6.3 no permitieron descartar en ningún caso el modelo de un único QTL.

Tabla 6.3. Resultado del análisis de dos QTLs para cada carácter. Donde Pos1 es la posición del QTL1, Pos2 la del QTL2, LR_0 es el máximo valor de LR resultado del contraste $M2a$ vs $M0$, P_0 el valor nominal de P correspondiente al citado análisis, LR_1 es resultado del contraste $M2a$ vs $M1a$, P_1 el valor nominal de P correspondiente a este análisis y a_1 y a_2 los efectos de los QTLs 1 y 2 respectivamente sobre el carácter.

Carácter	Pos ¹ .	Pos ² .	a_1 (ES)	a_2 (ES)	LR_0	P_0 -nominal	LR_1	P_1 - nominal
C18:2	33	61	-1.267 (0.395)	0.534 (0.410)	10.41	0.015	1.03	0.310
C20:4	31	50	-0.275 (0.226)	-0.259 (0.219)	10.88	0.012	1.33	0.249
C14:0	40	49	-0.074 (0.043)	0.137 (0.043)	13.36	0.004	3.05	0.081
C16:0	32	50	0.322 (0.298)	0.395 (0.293)	12.02	0.007	0.68	0.410

6.3.2. Análisis de detección de QTLs con 489 SNPs.

6.3.2.1. Modelo de un QTL

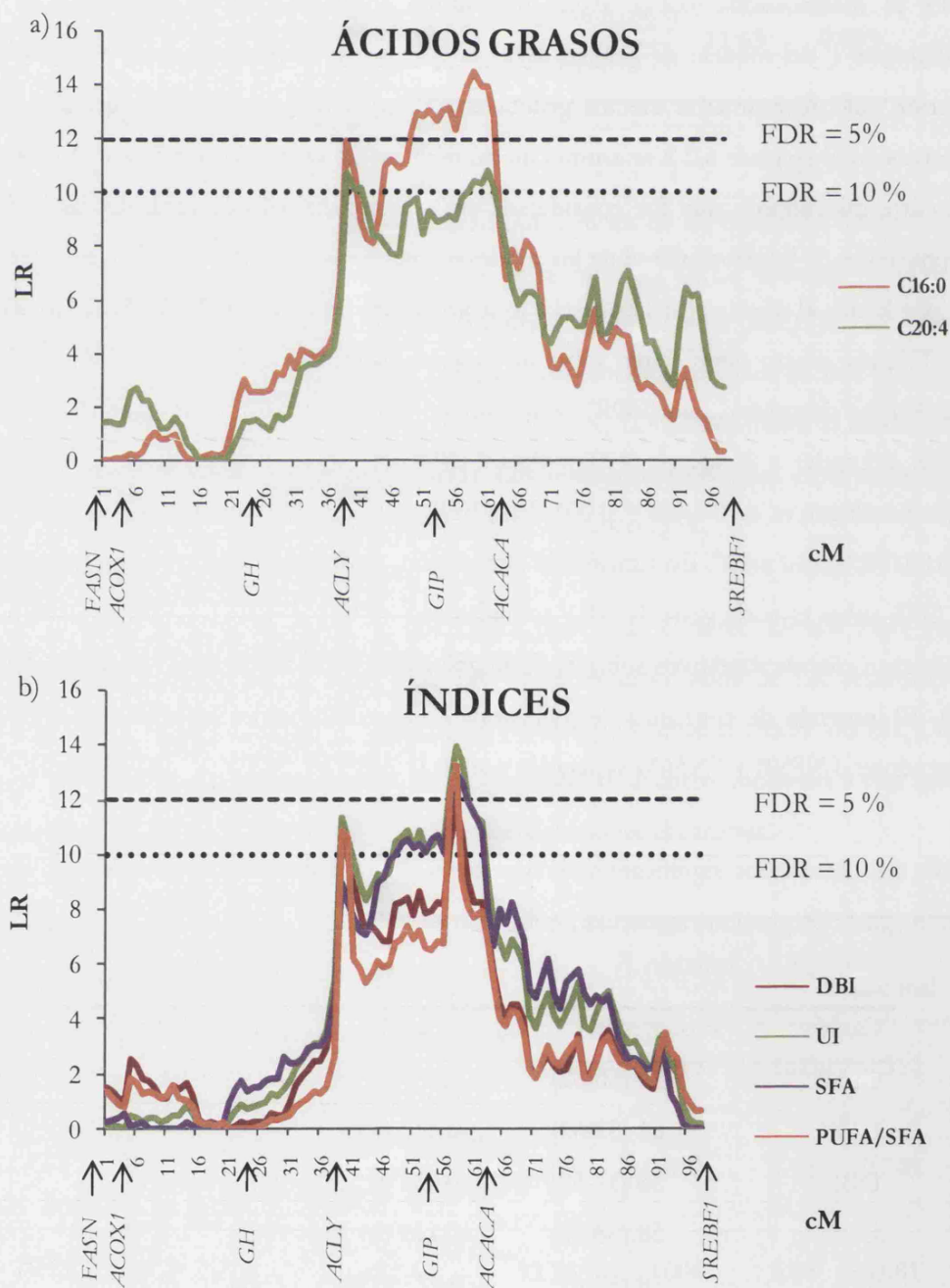
La detección de QTLs fue realizada a partir de los datos de genotipado de los 489 SNPs ya mencionados (6.2.2). Estos marcadores se utilizaron para el cálculo de los coeficientes c_i del modelo de análisis $M1a$, ya que, por el motivo expuesto en el análisis anterior, sólo se estimaron efectos genéticos aditivos. En función de la regresión lineal estimada en el epígrafe 6.2.2 se asumió un número medio de SNPs de 5.5 por cM con 0.2 cM entre marcadores, que fue considerada en la ejecución del programa QxPak 4.0. La comparación de la verosimilitud de los modelos con y sin QTL ($M1a$ y $M0$) se realizó cada cM por lo que el número total de tests a lo largo del cromosoma fue de 98. Para considerar un QTL con efecto significativo sobre un carácter realizando múltiples tests, se calculó la tasa de falso descubrimiento (FDR) de acuerdo con lo propuesto por Benjamini y Yekutieli (2005) y Lander y Kruglyak (1995). Para FDR=10% (QTL sugestivo) el umbral de P correspondiente es de $0.1/98 = 0.001$. Para FDR = 5% (QTL significativo), el valor de P es $0.05/98 = 5.10 \times 10^{-4}$. En función de este criterio, se detectaron efectos significativos de un QTL sobre el ácido graso C16:0 y sobre los índices UI, SFA y PUFA/SFA. Además se detectaron efectos sugestivos sobre el ácido graso C20:4 y el índice DBI (Tabla 6.4, Figura 6.4). El intervalo de confianza del QTL está comprendido entre 50 y 63 cM, entre los marcadores *SNP501* y *SNP631* separados por 10.11 Mb.

Tabla 6.4. Resultados significativos y sugestivos del análisis de detección de QTLs para ácidos grasos en grasa intramuscular, utilizando 489 SNPs.

Carácter	Posición en cM (IC)	a (ES)	LR	P-nominal
C16:0	59 (50-62)	0.667(0.174)	14.48 **	1.42×10^{-4}
C20:4	61 (53-63)	-0.443 (0.134)	10.86 *	9.82×10^{-4}
DBI	58 (57-59)	-0.033 (0.008)	11.73 *	6.13×10^{-4}
UI	58 (58-62)	-0.129 (0.034)	13.93 **	1.90×10^{-4}
SFA	58 (57-61)	0.827 (0.227)	12.94 **	3.21×10^{-4}
PUFA/SFA	58 (57-58)	-0.042 (0.014)	13.26 **	2.71×10^{-4}

* FDR = 10% ** FDR = 5%

Figura 6.4. Perfiles de los diferentes valores de la razón de verosimilitud (LR) del análisis de detección de QTLs basado en 489 SNPs localizados en SSC12. Sobre el eje de abscisas se indican las localizaciones teóricas de los genes candidatos. a) Ácidos grasos afectados por el QTL. b) Índices afectados por el QTL.



Los efectos sobre el ácido graso C16:0 y los índices UI y SFA corresponden, muy probablemente, al QTL2 detectado en el análisis de información combinada en la región próxima al gen *ACACA*. En esta misma región se detectaron también efectos sobre C20:4,

DBI, y PUFA/SFA a diferencia del análisis con información combinada en el que el valor máximo de dichos efectos se situó en una posición cercana al gen *ACLY*.

Por otro lado, los QTLs con efecto sobre los ácidos grasos C14:0 y C18:2 y los índices PI y PUFA detectados en el análisis de información combinada no superaron el umbral de $FDR < 10\%$ en el presente análisis. Los resultados relativos a estos caracteres se presentan en la tabla 6.5.

Tabla 6.5. Resultado del análisis de detección de un solo QTL, con 489 SNPs, en grasa intramuscular para aquellos caracteres en los que se obtuvieron resultados significativos en el análisis de información combinada.

Carácter	Posición en cM (IC)	a (ES)	LR	P -nominal
C14:0	59	0.064 (0.022)	9.38	0.0022
C18:2	59	-0.659 (0.303)	4.72	0.0297
PI	58	-2.562 (0.821)	9.67	0.0019
PUFA	58	-1.212 (0.435)	7.69	0.0055

6.3.2.2. Análisis multicarácter

Al igual que en el análisis de información combinada, el perfil del valor de verosimilitud y la posición en el mapa para la cual el valor de LR es máximo fueron similares en los ácidos grasos C14:0, C16:0, y C20:4. Para analizar si un modelo multicarácter de un QTL con efecto sobre estos tres ácidos grasos era más verosímil, se testaron tres modelos de 1 QTL con efecto sobre los dos ácidos grasos correspondientes con las tres parejas posibles que se pueden formar agrupando C14:0, C16:0, y C20:4.

Tabla 6.6. Resultados del análisis de un QTL multicarácter por parejas.

Parejas	Pos. (IC)	LR	P -nominal	Comparación
C14:0 - C16:0	58	14.51	7.03×10^{-4}	$14.51 - 14.48 = 0.03$
C14:0 - C20:4	59	14.05	8.86×10^{-4}	$14.05 - 10.86 = 3.19$
C16:0 - C20:4	59	16.05	3.27×10^{-4}	$16.05 - 14.48 = 1.57$

$\chi^2_{0.05[1]} = 3.84$ y $\chi^2_{0.01[1]} = 6.64$.

En la tabla 6.6 se muestran los resultados de estos tres análisis. Como resultado, al igual que en el análisis de información combinada, ninguno de ellos supera el umbral de significación correspondiente ($\chi^2_{0.01} = 3.84$) por lo tanto el modelo multicarácter no explicaría mejor los resultados. Esto sería consistente con el análisis anterior y consolidaría la hipótesis de que los efectos detectados sobre los ácidos grasos C14:0 y C20:4 son débiles y probablemente no significativos.

6.3.2.3. Modelos de dos QTLs

Como se ha comentado antes, al observar los perfiles de verosimilitud, se observa el máximo valor de LR localizado en la zona entre 50 y 63 cM coincidiendo con el QTL2 detectado en el análisis de información combinada, además de un segundo máximo localizado entre 36 y 45 cM, coincidiendo con el QTL1 detectado en el análisis de información combinada.

Para verificar la posible presencia de dos QTLs con efecto sobre los diferentes caracteres localizados en estas regiones, se realizó el contraste de la hipótesis nula de un QTL frente a dos. En primer lugar se testó un modelo de dos QTLs (M_{2a}) frente al modelo nulo (M_0) y posteriormente se testó frente a un modelo de un QTL (M_{1a}) (Tabla 6.7).

Los resultados de LR obtenidos al testar ambos modelos no superan el umbral de significación correspondiente ($\chi^2_{0.01} = 5.99$ para M_{2a} vs M_0 y $\chi^2_{0.01} = 3.84$ para M_{2a} vs M_1) para ninguno de los caracteres excepto para el ácido graso C14:0, pero hay que recalcar que en el análisis de un QTL éste no supera el umbral $FDR < 10\%$. Asimismo, en todos los casos, los efectos detectados son de signo contrario y particularmente, los máximos valores de LR del QTL con efecto sobre los ácidos grasos C14:0 y C16:0 están localizados muy próximos (43 y 58 cM).

Por lo tanto, a pesar de que los perfiles de verosimilitud parecen indicar la presencia de dos QTLs, los resultados obtenidos tanto en el análisis de información combinada como en el análisis de 489 marcadores apuntan que solo hay un QTL situado en la zona próxima a donde se localiza el gen *ACACA*, en distancia física 26.65 - 36.76 Mb, y los efectos detectados en la región del denominado como QTL1, localizado aproximadamente entre 17.34 - 22.26 Mb, estarían provocados por un fuerte desequilibrio de ligamiento en esa zona.

Tabla 6.7. Resultado del análisis de dos QTLs para cada carácter. Donde Pos1 es la posición del QTL1, Pos2 la del QTL2, LR_0 es el máximo valor de LR resultado en el contraste $M2a$ vs $M0$, P_0 el valor nominal de P correspondiente al citado análisis, LR_1 es resultado del contraste $M2a$ vs $M1a$, y P_1 el valor nominal de P correspondiente y a_1 y a_2 los efectos de los QTLs 1 y 2, respectivamente, sobre el carácter.

Carácter	Pos ¹ .	Pos ² .	a_1 (ES)	a_2 (ES)	LR_0	P_0 nominal	LR_1	P_1 nominal
C14:0	43	58	-0.198 (0.065)	0.253 (0.065)	18.23	3.93×10^{-4}	8.25	0.016
C16:0	43	58	-1.178 (0.538)	1.790 (0.539)	19.20	2.48×10^{-4}	4.53	0.104
C20:4	21	59	0.318 (0.173)	-0.625 (0.172)	12.65	0.005	0	-
DBI	21	58	0.024 (0.011)	-0.041 (0.011)	13.28	0.004	4.20	0.122
UI	21	58	0.096 (0.045)	-0.177 (0.044)	16.80	7.76×10^{-4}	4.09	0.129
SFA	28	58	-0.696 (0.349)	1.338 (0.343)	15.70	0.001	3.71	0.156
SFA/PUFA	25	58	1.575 (0.633)	-2.141 (0.624)	13.56	0.004	5.65	0.059

6.4. ANÁLISIS DE ASOCIACIÓN

6.4.1. Análisis de genes candidatos *ACLY*, *GIP* y *ACACA*

6.4.1.1. Modelo animal estándar

Pese a no estar incluidos en el mapa de ligamiento construido para realizar el barrido de QTLs se evaluaron como genes candidatos posicionales aquellos candidatos funcionales próximos a las dos regiones en donde se han localizado los QTLs 1 y 2.

En este sentido, los genes *ATP Citrate Lyase (ACLY)* y *Gastric Inhibitor Polipeptide (GIP)* fueron evaluados como candidatos para la región del QTL1 ya que se encuentran en

posiciones próximas a este intervalo. Para realizar el análisis de asociación con el modelo animal estándar se eligieron los polimorfismos *ACLY*: c.*523T>C y *GIP*: c.77A>G, este último está localizado en un exón y produce un cambio aminoacídico Asp26Gly aunque no tiene un efecto sobre la función de la proteína (Muñoz *et al.*, 2007). De la misma forma, el gen *Acetyl-CoA Carboxilase (ACACA)* fue evaluado como candidato para el QTL2. En este caso, se realizó el análisis de asociación con el polimorfismo *ACACA*: c.5634T>C. Los resultados obtenidos (Tabla 6.8) demuestran que no se detecta asociación significativa en ninguno de los casos.

Tabla 6.8. Resultados del análisis de asociación para los polimorfismos *ACLY*: c.*523T>C, *GIP*: c.77A>G y *ACACA*: c.5634T>C con el modelo animal estándar.

SNP	Carácter	LR	P	a_{SNP} (ES)
<i>ACLY</i> : c.*523T>C	C14:0	1.38	0.240	-0.021 (0.018)
	C16:0	1.95	0.163	-0.179 (0.128)
	DBI	0.94	0.333	0.008 (0.008)
	UI	1.65	0.198	0.036 (0.028)
	SFA	2.61	0.106	-0.275 (0.170)
	PUFA SFA/	0.93	0.336	0.013 (0.011)
<i>GIP</i> c.77A>G	C14:0	0.18	0.674	-0.014 (0.034)
	C16:0	0	-	-
	DBI	0	-	-
	UI	0.54	0.464	0.038 (0.052)
	SFA	1.28	0.258	-0.335 (0.296)
	PUFA/SFA	0.24	0.625	0.101 (0.022)
<i>ACACA</i> c.5634T>C	C14:0	1.11	0.292	0.019 (0.018)
	C16:0	1.45	0.227	0.171 (0.142)
	DBI	0.98	0.322	0.009 (0.009)
	UI	0.01	0.905	-0.006 (0.053)
	SFA	0.85	0.358	0.006 (0.006)
	PUFA/SFA	0.42	0.519	0.026 (0.040)

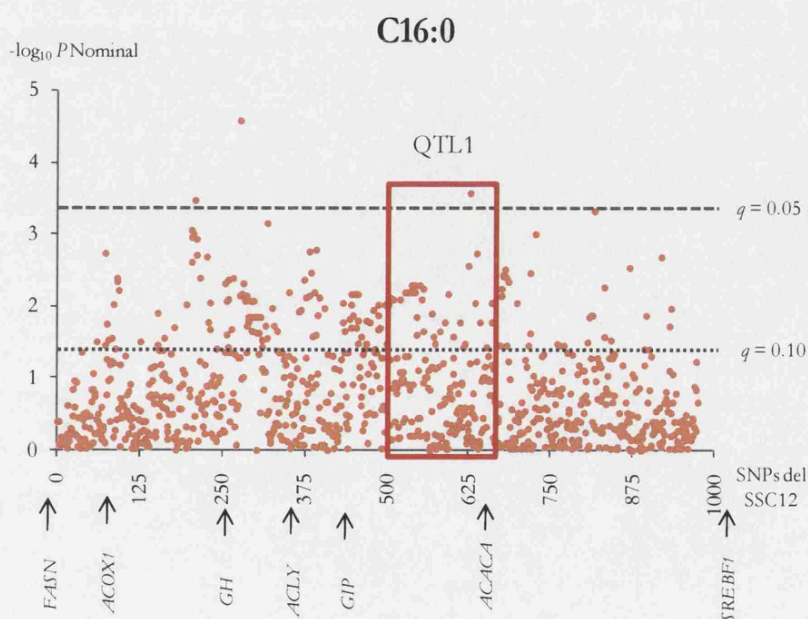
6.4.2 Análisis basado en SNPs del *chip* SNP60

6.4.2. Modelo animal estándar

Como en los casos anteriores, se realizó el análisis de asociación mediante un modelo animal estándar. Se aplicó un test de falso descubrimiento (FDR) para corregir el test múltiple para lo que se utilizó el paquete de R, *q-value* (Storey y Tibshirani, 2003), particularmente adecuado cuando se realiza un elevado número de contrastes (978). El umbral elegido para considerar significativo un marcador con efecto sobre determinado carácter fue en este caso el de un *q-value* < 0.10 .

Los caracteres analizados fueron los mismos que en los análisis anteriores de detección de QTLs. Se detectaron asociaciones significativas entre varios SNPs con el ácido palmítico (C16:0) y el cociente araquidónico/linoleico (C20:4/C18:2). En el caso del análisis del C16:0, se detectaron 167 SNPs con un valor de *q* inferior a 0.10, 98 de los cuales están localizados en regiones intergénicas y los 69 restantes están localizados en intrones de genes anotados (Figura 6.5).

Figura 6.5. \log_{10} del valor nominal de *P* resultante del análisis de asociación entre los 978 SNPs seleccionados y el carácter C16:0. En el eje de abscisas se representan las localizaciones teóricas de los genes candidatos. El recuadro rojo delimita la región dónde se detectó el QTL con efecto sobre este carácter en el barrido con 489 marcadores.



Cabe destacar que cuatro SNPs están fuertemente asociados con el carácter ya que su valor de q es inferior a 0.05 (Tabla 6.9). Uno de estos cuatro, el *SNP627* (sonda *DBWU0000115*), localizado en una región intergénica, se sitúa muy próximo a la posición más verosímil del QTL y tiene un efecto sobre el ácido graso C16:0 de 0.422 ± 0.128 . Los otros tres restantes no están próximos a ninguna de las zonas donde se detectaron QTLs y a pesar de localizarse en intrones de genes anotados, son genes que aparentemente no tienen una función relacionada con el carácter.

Tabla 6.9. Resultados del análisis de asociación entre los SNPs con números de orden 207, 208, 277 y 627 y C16:0, junto al SNP se muestra el nombre de la sonda entre paréntesis. Se indica la posición en Mb del SNP y el valor q (calculado por el programa *Q-value*)

SNP (sonda)	Mb	a_{SNP} (ES)	LR	P	q
<i>SNP207 (ASGA0053216)</i>	10.61	-0.493 (0.151)	10.91	9.54×10^{-4}	0.045
<i>SNP208 (ALGA0065055)</i>	10.63	-0.493 (0.151)	10.91	9.54×10^{-4}	0.045
<i>SNP277 (MARC0020580)</i>	22.96	-0.507 (0.134)	14.53	1.37×10^{-4}	0.014
<i>SNP627 (ASGA0054417)</i>	36.50	-0.422 (0.128)	11.10	8.62×10^{-4}	0.045

Por otro lado, dentro del grupo de SNPs con un valor de q comprendido entre 0.05 y 0.10, se destacan los SNPs 542, 544 y 545 (Tabla 6.10) localizados en el gen *Phosphatidylcholine Transfer Protein (PCIP)* que codifica para una proteína implicada en el transporte lipídico y que, a su vez, regula genes que codifican otras proteínas implicadas en el metabolismo de ácidos grasos como *PPARG*, *PPARA* y *CREBP*. Por lo tanto, este gen al estar localizado dentro del intervalo de confianza del QTL, podría ser un buen candidato funcional y posicional. Asimismo pertenecería a este grupo de sondas el *SNP636 (ALGA0066302)*, localizado en el gen *ACACA*.

Tabla 6.10. Resultados del análisis de asociación entre los SNP con números de orden 542, 544, 545 y 636 y C16:0.

SNP (sonda)	Mb	a_{SNP} (ES)	LR	P	q
<i>SNP542 (ASGA0054039)</i>	29.70	-4.320 (0.153)	7.81	0.005	0.055
<i>SNP544 (ASGA0054041)</i>	29.74	0.321 (0.128)	6.44	0.011	0.055
<i>SNP545 (ASGA0054044)</i>	29.82	0.350 (0.154)	5.25	0.022	0.086
<i>SNP636 (ALGA0066302)</i>	36.94	-0.400 (0.156)	6.75	0.009	0.055

En el análisis de asociación entre el índice C20:4/C18:2 y los 978 SNPs del SSC12, se detectaron 39 SNPs con un valor de q inferior a 0.10 (Figura 6.5). Ninguno de los SNPs tuvo un valor de q inferior a 0.05, siendo el más significativo el correspondiente al SNP881 (*H3GA0034898*) localizado en un intrón del gen *SCO* con un valor de q igual a 0.066 (Tabla 6.11). Este gen no codifica para una proteína con función clara relacionada con el metabolismo de ácidos grasos. De éstos, 27 están localizados en regiones intergénicas y los 12 restantes en genes anotados. El SNP845 (*H3GA0034702*) está localizado en el gen *Arachidonate 12-lipoxygenase (ALOX12)*. Este gen codifica para la proteína lipoxigenasa araquidonato 12-lipoxigenasa (12-LOX), que cataliza la oxidación del carbono 12 del araquidonato en ácido 12-hidroxicicosaenoico. Nótese que en los análisis de QTLs basados en la hipótesis de alelos alternativos en las líneas parentales, no se detectó anteriormente ninguno con efecto sobre este carácter.

Figura 6.5. Log_{10} del valor nominal de P resultante del análisis de asociación entre los 978 SNPs seleccionados y el carácter C20:4/C18:2.

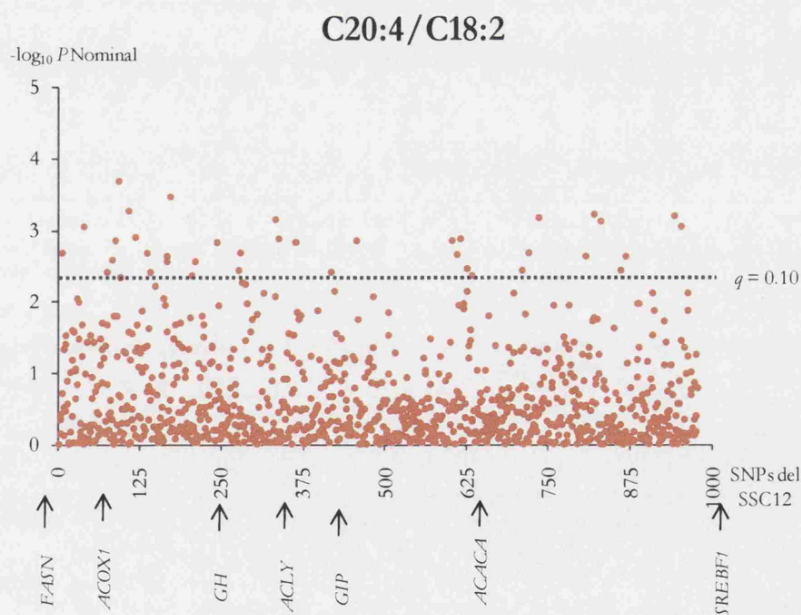


Tabla 6.11. Resultados del análisis de asociación entre los SNP con números de orden 881 y 845 y el índice C20:4/C18:0.

SNP (sonda)	Mb	a_{SNP} (ES)	LR	P	q
SNP881 (<i>ASGA0054039</i>)	52.47	-0.024 (0.007)	13.77	2.07×10^{-4}	0.066
SNP845 (<i>ASGA0054041</i>)	48.67	0.017 (0.007)	6.73	0.009	0.094

Además, en el caso del ácido graso C20:4 y los índices PI y PUFA/SFA se detectaron SNPs con valores de q muy próximos al nivel de significación (Figura 6.6). En primer lugar, se detectaron 46 SNPs con un valor de q próximo a 0.11 para el ácido graso C20:4, de los cuales, 34 están localizados en regiones intergénicas y 12 en regiones intrónicas de genes conocidos. Entre ellos se incluyen los SNPs 542 y 544 que también tienen efecto sobre el C16:0 y como se indicaba anteriormente están situados en el gen *PCTP*.

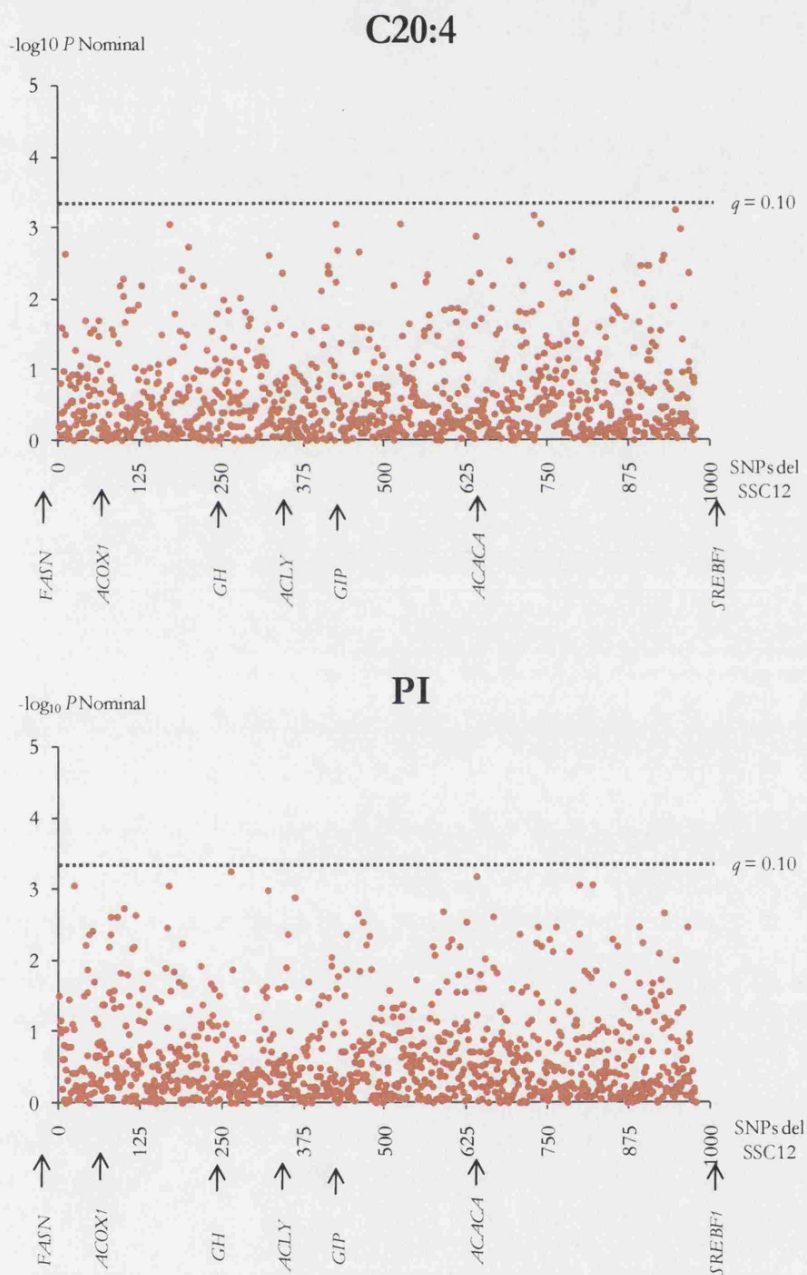
Se detectaron otros 15 SNPs con efecto sobre el índice PI con un valor $q < 0.12$, de los cuales 10 están situados en regiones intergénicas y cinco en intrones de genes anotados. Uno de estos SNPs es el ya mencionado SNP 544 situado en el gen *PCTP*.

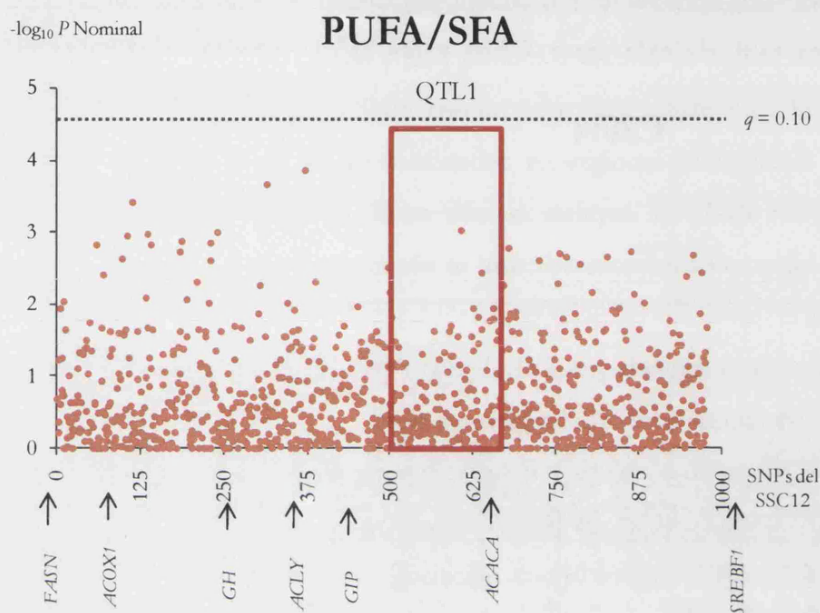
Por último, se detectaron dos SNPs con un valor de $q < 0.10$ y 19 con un valor inferior a 0.12 en el análisis de asociación con el índice PUFA/SFA. Del total, 14 son intergénicos y siete intrónicos, uno de los cuales es el ya mencionado SNP 544. Los resultados del análisis de asociación referente a los SNPs 542 y 544 con efecto sobre estos tres ácidos grasos se detallan en la tabla 6.12.

Tabla 6.12. Resultados del análisis de asociación con el modelo animal estándar para el SNP542 con efecto sobre C20:4 y el SNP544 con efecto sobre el ácido graso C20:4 y los índices PI y PUFA/SFA.

SNP (sonda)	Mb	Caracter	a_{SNP} (ES)	LR	P	q
SNP542 (<i>ASGA0054039</i>)	29.70	C20:4	-0.303 (0.118)	6.19	0.009	0.102
		C20:4	-0.264 (0.097)	7.91	0.004	0.102
SNP544 (<i>ASGA0054041</i>)	29.74	PI	-1.665 (0.603)	8.15	0.004	0.115
		PUFA/SFA	-0.024 (0.009)	9.68	0.002	0.115

Figura 6.6. Log_{10} del valor nominal de P resultante del análisis de asociación entre los 978 SNPs seleccionados y a) el ácido graso C20:4 y los índices b) PI y c) PUFA/SFA.





6.4.3. MAAT.

Debido a que se detectó un efecto significativo entre el *SNP636* (*ALGA0066302*) localizado en el gen *ACACA*, se efectuó un análisis de asociación asistida por marcadores. Cuando se contrastó el efecto del polimorfismo incluyendo el efecto del QTL, los efectos del primero desaparecieron y cuando se realizó la segunda parte del test contrastando los efectos del QTL incluyendo los efectos del SNP, los efectos del QTL se mantuvieron. Por lo tanto esto indicaría que la mutación causal tendría un desequilibrio de ligamiento mayor con el QTL que con el SNP (Tabla 6.13).

Tabla 6.13. Resultado del test de asociación asistida por marcadores entre el polimorfismo *SNP636* y C16:0.

Efecto del SNP					QTL		
SNP	a (ES)	LR	P-value	cM	a (ES)	LR	P-nominal
<i>SNP636</i>	-0.125 (0.142)	0.80	0.371	49	0.674 (0.177)	14.16	0.0017

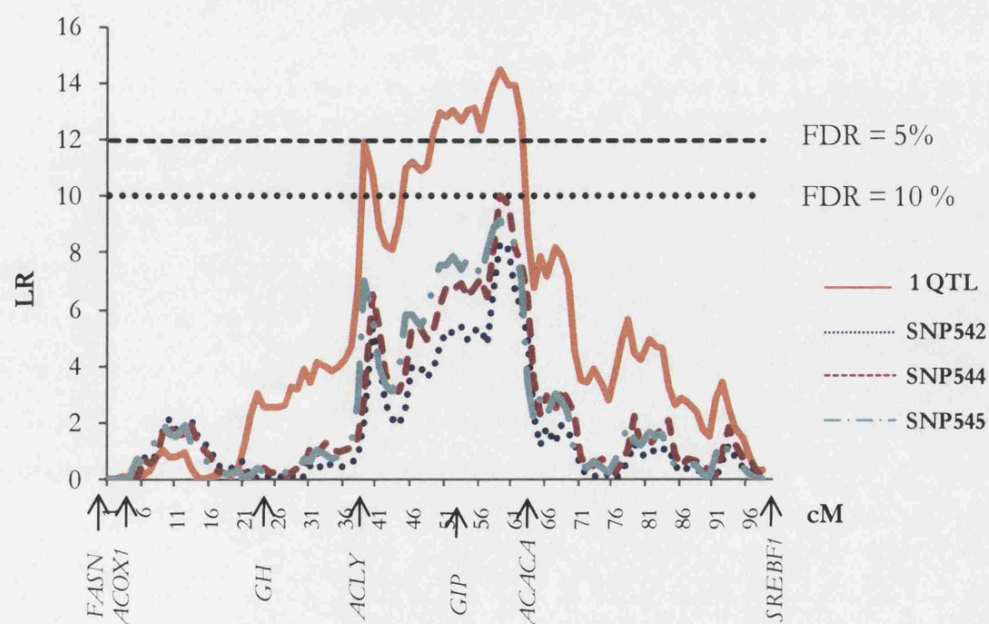
Además se testaron mediante este modelo los SNPs 542, 544 y 545. Como resultado, los efectos de los SNPs se mantuvieron y el máximo valor del perfil de verosimilitud disminuyó en todos los casos. Esto indicaría que la mutación causal se

encontraría en fuerte desequilibrio de ligamiento con los SNPs del gen PCTP (Tabla 6.14 y Figura 6.7).

Tabla 6.14. Resultado del test de asociación asistida por marcadores entre los polimorfismos *SNP542*, *SNP544* y *SNP545* y C16:0.

SNP	Efecto del SNP				QTL		
	a (ES)	LR	P-value	cM	a (ES)	LR	P-nominal
<i>SNP542</i>	-0.265 (0.182)	9.08	0.003	58	0.685 (0.232)	8.40	0.004
<i>SNP544</i>	0.349 (0.129)	7.48	0.006	58	0.699 (0.215)	10.03	0.001
<i>SNP545</i>	0.398 (0.156)	6.28	0.012	58	0.745 (0.244)	9.18	0.002

Figura 6.7. Perfil de la razón de verosimilitud de QTL cuando se incluye en el modelo el polimorfismo correspondiente. Las líneas continuas representan el análisis de QTL para C16:0 con el modelo de un QTL y las discontinuas el cambio de significación producido al incluir el SNP indicado



6.5. ANÁLISIS DEL POLIMORFISMO *ACACA*: c.5634T>C EN CERDOS IBÉRICOS PUROS Y CRUZADOS CON DUROC

Paralelamente al estudio realizado en grasa subcutánea, se analizó el efecto del polimorfismo *ACACA*: c.5634T>C en el perfil de ácidos grasos de grasa intramuscular, ya que se disponía de estas medidas en los dos ensayos con animales comerciales (Tablas 3.3 y 3.4).

6.5.1. Animales *Ibérico x Ibérico*

Se genotiparon un total de 108 animales pertenecientes al núcleo de selección de Ibéricos puros con registros de ácidos grasos recogidos en grasa intramuscular. Se observaron frecuencias alélicas similares a las obtenidas en el estudio de grasa subcutánea (T = 0.69 y C = 0.31).

Se detectó asociación significativa entre el ácido graso C20:1(n-9) y el polimorfismo con un valor máximo de LR igual a 5.18 ($P < 0.023$), obteniéndose un efecto aditivo de la sustitución del alelo T por el C de 0.51 ± 0.02 , a diferencia del análisis de asociación realizado en el cruce IBMAP entre este polimorfismo y el perfil de ácidos grasos de grasa intramuscular en el que no se detectaron efectos significativos.

6.5.2. Animales *Duroc x Ibérico*

Se genotiparon para el citado polimorfismo los mismos animales de esta población que se emplearon en el estudio de grasa subcutánea.

Se detectó asociación significativa entre el SNP *ACACA*: c.5634T>C y el perfil de ácidos grasos de grasa intramuscular recogido en la tabla 6.15. Cabe destacar que el alelo C tiene un efecto negativo en los porcentajes de ácidos grasos saturados, y positivo sobre los de ácidos grasos insaturados. Como se mencionaba anteriormente, en el cruce IBMAP no se detectó ningún efecto significativo del polimorfismo sobre ninguno de los caracteres analizados.

Tabla 6.15. Resultados del análisis de asociación en grasa intramuscular con resultados significativos entre el polimorfismo *ACACA*: c.5634T>C y los ácidos grasos C14:0 y C16:0.

Carácter	LR	P	a_{SNP} (ES)
C16:0	4.23	0.040	-0.499 (0.238)
C17:0	5.33	0.021	0.006 (0.002)
C17:1	5.91	0.015	0.010 (0.004)
C18:0	3.60	0.058	-0.375 (0.192)
C18:2 (n-3)	4.99	0.026	0.236 (0.101)
C18:2 (n-6)	4.11	0.043	0.011 (0.005)
C22:5 (n-3)	4.02	0.045	0.005 (0.003)
SFA	4.17	0.041	-0.887 (0.423)
PUFA	4.77	0.029	0.302 (0.133)

*Análisis
de expresión*

7.1. ANÁLISIS ESTADÍSTICO DE DATOS DE EXPRESIÓN

La implementación de estos microarrays estuvo enfocado, en un principio, a la caracterización a nivel genómico de patrones de expresión diferencial y redes genéticas implicadas en crecimiento porcino y caracteres de calidad de carne., según se recoge en el subproyecto AGL2008-04818-C03-02 del Plan Nacional de Investigación (2008-2011) llevado a cabo por el Área de Producción Animal del Centre UdL-IRTA (Lleida, España). Por lo que en este centro se desarrollaron las tareas de extracción de ARN y posteriores transcripciones in vitro del mismo, tal y como se recoge en el epígrafe 3.7.1 de la sección Materiales y Métodos.

Se llevaron a cabo diferentes análisis de expresión con el objetivo de encontrar diferencias de expresión en genes localizados en el SSC12 y que codificasen para alguna proteína con una implicada en el metabolismo de ácidos grasos.

Se analizó la expresión mediante microarrays de 100 individuos pertenecientes a la generación RC2 del cruce IBMAP. Los 100 microarrays pasaron todos los análisis de calidad por lo que eran comparables entre sí. Los microarrays fueron normalizados mediante el método GCRMA y una vez normalizados se realizó el análisis estadístico de los mismos. Para realizar el análisis se utilizó el programa GEAMM v.1.6. En el modelo mixto fueron incluidos como efectos fijos:

- Media. Efecto discreto compuesto de un nivel.
- Sexo. Efecto discreto compuesto por dos niveles: macho y hembra.
- Lote de sacrificio. Efecto discreto compuesto por tres niveles dependiendo del lote en el que el animal fue sacrificado; lote 1, lote 2 y lote 3.
- Peso de la canal al sacrificio. Efecto continuo que tiene una distribución normal
- Efecto problema. Se realizaron tres análisis distintos en los que cambiaba el efecto problema. Este era discreto o continuo en función de la clasificación empleada para analizar las diferencias de expresión entre los individuos. Las clasificaciones fueron realizadas en función de los resultados obtenidos en el estudio de composición de ácidos grasos de grasa intramuscular:

- ✓ Clasificación 1. Los individuos fueron clasificados en dos grupos en función de la probabilidad de ser *Qq* o *qq* para el QTL detectado en el análisis con 489 marcadores cuyo valor máximo de LR está situado en 59 cM (31.89 Mb) y está acotado por los marcadores *SNP501* y *SNP631* (26.65 - 36.76 Mb). El efecto fue

introducido en el modelo como discreto con dos niveles; 1 para una probabilidad superior a 0.9 de ser *Qq* y 2 si tenía una probabilidad superior a 0.9 de ser *qq*. En total fueron incluidos en este análisis un total de 99 individuos ya que uno de ellos tiene una probabilidad de origen inferior a 0.9

✓ Clasificación 2. Los individuos fueron clasificados en función del genotipo (GG, GA, AA) del *SNP636* (*ASGA0066302G>A*) del *chip SNP60*. Ya que en el análisis de asociación con los datos del *chip* se detectó asociación significativa sobre el ácido graso C16:0 y además, éste se localiza dentro del gen *ACACA*. El efecto fue introducido como una covariable de valor 0 si el genotipo del individuo no tenía ninguna copia del alelo A, 1, si el genotipo sólo tenía una copia y 2 si tenía dos copias.

✓ Clasificación 3. Los individuos fueron clasificados en función del genotipo (TT, CT, CC) del *SNP542* (*ASGA0054039T>C*) del *chip SNP60*, que está localizado dentro del gen *PCTP* y para el que también se detectó asociación significativa sobre el ácido graso C16:0. El efecto fue introducido como covariable de valor 0 si el genotipo del individuo no tenía ninguna copia del alelo C, 1, si sólo tenía una copia y 2 si tenía dos copias.

✓ Clasificación 4. Los individuos fueron clasificados en función del genotipo (AA, AG, GG) del *SNP544* (*ASGA0054041A>G*) del *chip SNP60*. Este SNP está también localizado dentro del gen *PCTP* y además también se obtuvieron resultados significativos en el análisis con C16:0 así como una asociación sugestiva sobre el ácido graso C20:4 y los índices PI y PUFA/SFA. El efecto fue introducido como covariable de valor 0 si el genotipo del individuo no tenía ninguna copia del alelo G, 1 si tenía una copia y 2, para dos copias.

✓ Clasificación 5. La clasificación fue similar a la clasificación 3 excepto que el efecto era introducido como discreto realizándose las comparaciones entre las diferentes parejas formadas.

En el análisis se consideró como significativa la expresión de aquellas sondas que presenten un $p < 0.001$ y además que tuviesen un ratio de diferencia de 1.10, nótese que el ratio en la escala no transformada de un efecto discreto es calculado como 2^δ (siendo δ la

media posterior de la diferencia entre dos niveles de un efecto discreto) o $2^{S \times \beta}$ (siendo S la desviación estándar de un efecto continuo y β la media posterior de su coeficiente de regresión. En todos los casos se aplicó un FDR del 10%.

7.1.1. Resultados en función de la clasificación 1: Prob (Qq)/Prob (qq)

Mediante este análisis se detectaron 22 sondas con $p < 0.001$, que al aplicar un FDR del 10% quedaron reducidas a cuatro. Estas cuatro sondas se representan en color rojo en la tabla 7.1.

Tabla 7.1. Genes diferencialmente expresados entre los grupos con genotipos probables para el QTL Qq y qq .

Sonda	Anotación	Cromosoma	Mb	$P_{\text{posterior}}$	Ratio
Ssc.24800.2.A1_at	<i>LOC100514982</i>	SSC12	24.43	1.10×10^{-6}	1.21
Ssc.18579.1.A1_at	<i>CDK5RAP3</i>	SSC12	21.68	5.00×10^{-6}	1.31
Ssc.9883.1.A1_at	<i>ADRBK2</i>	SSC14	43.38	8.07×10^{-6}	0.74
Ssc.6484.1.S1_at	<i>PNPO</i>	SSC12	21.64	2.56×10^{-5}	0.70
Ssc.3964.1.S1_at	<i>CUEDC1</i>	SSC12	-	4.30×10^{-5}	1.36
Ssc.877.1.S1_at	<i>SKA2</i>	SSC12	-	4.41×10^{-5}	1.24
Ssc.22988.1.S1_at	<i>PCTP</i>	SSC12	29.65	9.93×10^{-5}	0.77
Ssc.30925.1.A1_at	<i>DKK2</i>	SSC8	98.74	0.00016	0.74
Ssc.1286.1.A1_at	<i>SUZ12</i>	SSC12	40.82	0.00016	1.21
Ssc.24930.1.S1_at	<i>LTV1</i>	SSC1	22.43	0.00020	1.18
Ssc.13777.2.S1_at	<i>TMEM97</i>	SSC12	26.64	0.00027	1.26
Ssc.21326.1.S1_at	<i>NEO1</i>	SSC7	65.55	0.00035	1.23
Ssc.2318.1.A1_at	<i>GLI2</i>	SSC15	28.58	0.00038	0.87
Ssc.22443.2.A1_at	<i>TSR1</i>	SSC12	45.85	0.00038	1.13
Ssc.7149.3.S1_at	<i>CYB5D1</i>	SSC12	-	0.00046	1.15
Ssc.11077.1.S1_at	<i>POU2F1</i>	SSC4	86.90	0.00050	1.24
Ssc.12255.1.A1_at	<i>FSTL5</i>	SSC8	43.84	0.00061	1.15
Ssc.7624.1.S1_at	-	-	-	0.00067	1.13
Ssc.24711.1.S1_at	<i>SMYD4</i>	SSC12	45.37	0.00070	1.15
Ssc.428.5.S1_at	-	-	-	0.00074	0.83
Ssc.29805.1.A1_at	<i>SKA2</i>	SSC12	-	0.00083	1.22
Ssc.14311.1.A1_at	<i>KLAA1407</i>	SSC13	10.90	0.00087	1.18

Se anotaron todas las sondas diferencialmente expresadas excepto dos. Para realizar esta anotación, además de la proporcionada por Tsai *et al.*, (2006) se realizaron diferentes alineamientos entre las secuencias de las sondas del *microarray* y las bases de datos disponibles en la red.

Una vez anotados las sondas, se utilizó la herramienta GO (DAVID 2008) para analizar si los términos GO-BP (funciones implicadas en procesos biológicos) estaban relacionados con el metabolismo de ácidos grasos. La herramienta agrupa los procesos e indica que sólo pueden considerarse como significativos aquellos que tengan una puntuación superior a 1.5. Ninguno de los grupos superó el umbral de 1.5, esto se puede deber a que las proteínas codificadas por los genes diferencialmente expresados no tengan funciones similares y no se puedan agrupar en grupos funcionales, o bien que el número de genes diferencialmente expresados es muy bajo y el programa no tiene suficiente información para realizar los grupos.

Se localizó la posición de los genes en los diferentes cromosomas porcinos mediante el uso de las bases de datos Ensembl y NCBI y cuando no fue posible porque el gen no estuviese mapeado, se recurrió al mapa comparativo humano-porcino, aunque en estos no ha sido posible determinar la posición exacta en Megabases.

De los 22 genes diferencialmente expresados, 11 están localizados en el SSC12. Cabe destacar que cinco de ellos se encuentran localizados entre las posiciones de 21.64 y 29.65 Mb, próximos al QTL o incluso incluidos dentro del intervalo de confianza del mismo (26.47-35.11 Mb). Uno de estos genes es el ya conocido *PCTP*. La proteína codificada por este gen pertenece a la familia STARD, de hecho este gen también es conocido como *STARD2*. El producto proteico está implicado entre otros en el transporte lipídico y de fosfolípidos y en el metabolismo del colesterol. Por lo tanto, en función de estos resultados y de los resultados obtenidos en el epígrafe 6.4.2.1, este gen es un sólido candidato funcional y posicional. Otro de los genes diferencialmente expresados que resulta interesante desde el punto de vista de su función y su localización es el gen *TMEM97* que codifica para una proteína de membrana que juega un papel importante controlando los niveles celulares de plasma y que cuya expresión está regulada por los factores de transcripción *SREBF1* y *SREBF2*. Los restantes diez genes no tenían una función directa relacionada con el metabolismo lipídico.

7.1.2. Resultados en función del número de copias del alelo *ACACA*: *ASGA0066302A*

Se detectaron 20 sondas con $p < 0.001$ (Tabla 7.2), todas ellas tenían un valor de $FDR < 0.10$. Todas las sondas pudieron ser anotadas. Asimismo mediante GO (DAVID 2008) se anotaron funcionalmente los 19 genes diferencialmente expresados. Sin embargo no se detectaron grupos funcionales con una puntuación superior a 1.5.

Tabla 7.2. Genes diferencialmente expresados en función del genotipo del número de copias de los alelos del *SNP636*

Sonda	Anotación	Cromosoma	Mb	$P_{\text{posterior}}$	Ratio*
Ssc.2881.1.S1_at	<i>CPSF3L</i>	SSC6	42.91	2.98×10^{-6}	1.12
Ssc.24636.1.A1_at	<i>PANK3</i>	SSC16	52.66	3.94×10^{-6}	1.13
Ssc.9782.1.A1_at	<i>KRT10</i>	SSC12	-	6.82×10^{-6}	1.19
Ssc.29848.1.A1_at	<i>PDE4D</i>	SSC16	35.86	2.76×10^{-5}	1.16
Ssc.3709.1.S1_at	<i>MRPL10</i>	SSC12	21.58	2.81×10^{-5}	1.17
Ssc.1286.1.A1_at	<i>SUZ12</i>	SSC12	40.82	0.00024	0.86
Ssc.19416.1.A1_at	<i>IL21R</i>	SSC3	17.6	0.00027	0.67
Ssc.23484.1.A1_a_at	<i>SCARB2</i>	SSC8	61.27	0.00028	0.86
Ssc.26888.1.A1_at	<i>DDX25</i>	SSC9	-	0.00029	1.11
Ssc.29640.1.A1_at	<i>RBM39</i>	SSC17	41.2	0.00036	1.12
Ssc.19050.1.S1_at	<i>DHRS7C</i>	SSC12	51.7	0.00038	1.21
Ssc.4026.1.S1_at	<i>ERVWE1</i>	SSC9	-	0.00047	1.12
Ssc.16926.1.S1_at	<i>SLC2A5</i>	SSC6	-	0.00051	1.11
Ssc.24698.1.S1_at	<i>GCHFR</i>	SSC1	137.06	0.00071	1.17
Ssc.9351.1.A1_at	<i>OSR1</i>	SSC13	18.95	0.00081	1.11
Ssc.4246.3.S1_at	<i>ARHGEF10L</i>	SSC6	52.46	0.00085	1.23
Ssc.15565.1.S1_at	<i>LCP2</i>	SSC16	50.98	0.00090	1.16
Ssc.14311.1.A1_at	<i>KLAA1407</i>	SSC13	10.89	0.00096	0.87
Ssc.8528.2.S1_at	<i>CTSC</i>	SSC9	-	0.00097	1.31
Ssc.8528.1.A1_at	<i>CTSC</i>	SSC9	-	0.00099	1.27

* Efecto de una desviación típica ($\sigma = 0.58$) del número de copias sobre la expresión de cada sonda

De las 20 sondas anotadas, cuatro corresponden a genes localizados en el SSC12. De estos cuatro, uno de ellos, el gen *MRPL10* está localizado en la posición de 21.58 Mb, cercano al QTL. No obstante, la proteína codificada por este gen no tiene efecto directo sobre la composición de ácidos grasos. Asimismo, el resto de los genes no tiene una función directamente relacionada con el metabolismo de ácidos grasos, por lo que ninguno ha podido ser considerado como un gen candidato potencial.

7.1.3. Resultados en función del número de copias del alelo *PCTP* *ASGA0054039C*

Se detectaron un total de 48 sondas diferencialmente expresadas con $p < 0.001$, de las cuales 8 tenían un valor de FDR < 0.10 (Tabla 7.3).

Tabla 7.3. Genes diferencialmente expresados en función del número de copias de los alelos del *SNP542*

Sonda	Anotación	Cromosoma	Mb	$P_{\text{posterior}}$	Ratio*
Ssc.22988.1.S1_at	<i>PCTP</i>	SSC12	29.65	3.34×10^{-11}	0.82
Ssc.5902.1.A1_at	<i>COX11</i>	SSC12	28.93	1.16×10^{-8}	1.20
Ssc.17449.1.S1_at	<i>OLFM1</i>	SSC1	-	1.63×10^{-8}	1.36
Ssc.10192.1.A1_at	<i>COX11</i>	SSC12	28.93	1.53×10^{-7}	1.13
Ssc.6525.2.A1_at	<i>PPP1R12A</i>	SSC5	96.06	1.07×10^{-5}	1.27
Ssc.26209.1.S1_at	<i>SPATS2</i>	SSC15	97.65	2.14×10^{-5}	0.89
Ssc.29609.2.A1_at	<i>SVIP</i>	SSC2	-	2.55×10^{-5}	1.27
Ssc.19050.1.S1_at	<i>DHRS7C</i>	SSC12	51.7	0.00006	0.88

* Efecto de una desviación típica ($\sigma = 0.51$) del número de copias sobre la expresión de cada sonda

Las 48 sondas diferencialmente expresadas pudieron ser anotadas, cabe mencionar que dos de ellas correspondían a distintos transcritos del mismo gen (*COX11*). Mediante la herramienta DAVID 2008 se anotaron funcionalmente 41 de los genes y tal como en el análisis anterior se intentó de agrupar de acuerdo a sus funciones biológicas. Ninguno de los grupos superó la puntuación de 1.5 indicada por el programa, probablemente debido a que no había información suficiente para realizar este análisis.

De los siete genes con FDR < 0.10 , *COX11*, *PCTP* y *DHRS7C* están localizados en el SSC12, los dos primeros en posiciones cercanas al intervalo de confianza del QTL (28.93 y 29.65 Mb) y el tercero, *DHRS7C* más alejado del intervalo de confianza del QTL (51.7 Mb). Cabe destacar que el gen diferencialmente expresado que presenta mayor significación estadística es el propio gen *PCTP*.

7.1.4. Resultados en función del alelo *PCTP*: *ASGA0054041G*

Se detectaron un total de 38 sondas diferencialmente expresadas con $p < 0.001$, todas ellas tenían un valor de FDR < 0.10 (Tabla 7.4).

Tabla 7.4. Genes diferencialmente expresados en función del número de copias de los alelos del del *SNP544*.

Sonda	Anotación	Cromosoma	Mb	<i>P</i> posterior	Ratio
Ssc.22988.1.S1_at	<i>PCTP</i>	SSC12	29.65	4.25×10^{-14}	1.34
Ssc.5902.1.A1_at	<i>COX11</i>	SSC12	28.93	7.4×10^{-9}	0.78
Ssc.7209.1.A1_at	<i>COX11</i>	SSC12	28.93	1.15×10^{-6}	1.11
Ssc.24800.2.A1_at	<i>LOC100514982</i>	SSC12	24.44	3.02×10^{-6}	0.89
Ssc.6484.1.S1_at	<i>PNPO</i>	SSC12	21.64	2.33×10^{-5}	1.27
Ssc.1023.1.S1_at	<i>RBM35B</i>	SSC6	-	3.21×10^{-5}	1.21
Ssc.30718.1.A1_at	<i>KCNH2</i>	SSC18	53.61	4.98×10^{-5}	1.30
Ssc.9577.1.S1_at	<i>ATXN7</i>	SSC13	38.26	6.64×10^{-5}	1.13
Ssc.10391.1.A1_at	<i>MCAM</i>	SSC9	45.13	7.34×10^{-5}	0.85
Ssc.19050.1.S1_at	<i>DHRS7C</i>	SSC12	51.70	7.48×10^{-5}	1.19
Ssc.17447.1.A1_at	<i>GTF3C5</i>	SSC1	288.40	8.17×10^{-5}	1.10
Ssc.2847.1.S1_at	<i>ANKRD9</i>	SSC7	131.30	0.00010	1.25
Ssc.24715.1.S1_at	<i>FAM8A1</i>	SSC7	139.19	0.00015	1.11
Ssc.4110.1.S1_at	<i>TTLA4</i>	SSC15	113.90	0.00017	1.14
Ssc.21245.1.S1_at	<i>GRAMD4</i>	SSC5	-	0.00018	1.11
Ssc.12654.1.A1_at	<i>MEF2A</i>	SSC1	146.05	0.00027	1.11
Ssc.25859.1.S1_at	<i>ANKRD40</i>	SSC12	23.98	0.00029	1.16
Ssc.12654.2.A1_at	<i>MEF2A</i>	SSC1	146.05	0.00030	1.13
Ssc.6978.1.A1_at	<i>MAX</i>	SSC1	136.03	0.00030	1.16
Ssc.6382.1.A1_at	<i>PPP1R3B</i>	SSC15	53.09	0.00031	1.37
Ssc.19884.1.S1_at	<i>SAV1</i>	SSC1	18.87	0.00047	1.14
Ssc.10549.1.A1_at	<i>BOC</i>	SSC13	109.37	0.00051	1.18
Ssc.15352.1.S1_at	<i>FAM117A</i>	SSC12	23.29	0.00054	1.14
Ssc.21383.1.A1_at	<i>PHOSPHO1</i>	SSC12	-	0.00063	1.15
Ssc.5620.1.A1_at	<i>TOB1</i>	SSC12	23.90	0.00065	1.12
Ssc.16679.1.S1_at	<i>BMP1</i>	SSC14	64.00	0.00069	0.78
Ssc.13476.1.A1_at	<i>PEG10</i>	SSC9	-	0.00074	0.75
Ssc.4895.2.S1_at	<i>FBXO25</i>	SSC15	-	0.00075	1.16
Ssc.10675.1.A1_at	<i>LNK1</i>	SSC8	34.63	0.00085	1.162
Ssc.29510.1.A1_at	<i>ZNF780A</i>	SSC6	-	0.00090	1.17
Ssc.26209.1.S1_at	<i>SPATS2</i>	SSC15	97.65	0.00090	1.14
Ssc.1333.1.A1_at	<i>ZCCHC11</i>	SSC6	113.45	0.00091	1.11
Ssc.8295.1.A1_at	<i>AGXT2L1</i>	SSC8	97.37	0.00094	1.36
Ssc.27429.1.A1_at	<i>CPEB2</i>	SSC8	74.72	0.00094	0.81
Ssc.2230.1.A1_at	<i>NOC3L</i>	SSC14	112.00	0.00096	0.90
Ssc.22197.1.S1_at	<i>FANCG</i>	SSC1	247.02	0.00097	0.84
Ssc.19381.1.A1_at	<i>HOMER3</i>	SSC2	-	0.00098	1.12

* Efecto de una desviación típica ($\sigma = 0.66$) del número de copias sobre la expresión de cada sonda

Todas las sondas pudieron ser anotadas. Asimismo mediante GO (DAVID 2008) se anotaron funcionalmente 33 de los 38 genes diferencialmente expresados. Sin embargo no se detectaron grupos funcionales con una puntuación superior a 1.5.

Un total de nueve genes fue localizado en el SSC12, siete de los cuales están localizados en la región 21.64 – 29.65 Mb, próximos al intervalo de confianza del QTL (30.96-35.02 Mb). La sonda diferencialmente expresada con mayor nivel de significación corresponde al propio gen *PCTP*. El valor del ratio igual a 1.31 indicaba un incremento en la expresión de un 31% producido por una desviación típica ($\sigma = 0.66$ copias), por lo tanto, una copia del alelo G incrementa la expresión en un 50% y dos la duplicaría.

7.1.5. Resultados en función de los genotipos del SNP *ASGA0054041A>G*

En este análisis, se analizó la expresión diferencial atendiendo a las comparaciones entre las parejas de los tres posibles genotipos, comparándose AA vs AG, AA vs GG y AG vs GG. Al comparar los genotipos de los dos homocigotos se detectaron un total de 57 genes diferencialmente expresados. En la comparación AA vs AG se detectaron 18 sondas diferencialmente expresadas, mientras que al comparar GG vs AG, se detectaron un total de 41 sondas diferencialmente expresadas. En la tabla 7.5 se muestran los ratios de las sondas diferencialmente expresadas al comparar los diferentes genotipos. En esta tabla sólo aparecían representadas aquellas que estaban diferencialmente expresadas en el análisis realizado utilizando la clasificación 4.

A partir de los resultados obtenidos en este análisis podemos saber cuál es la diferencia de expresión en cada sonda para cada uno de los genotipos. En el caso del gen *PCTP*, los individuos con el genotipo GG expresan este transcrito dos veces más que los homocigotos AA y un 30% más que los heterocigotos. Asimismo los heterocigotos expresan el transcrito un 30% más que los homocigotos AA. Estos resultados indican que el efecto del SNP sobre la expresión del gen es aditivo.

Los genes *COX11*, *PNPO*, *ANKRD40*, *FAMM117A*, *LOC100514982*, *CDK5RAP3*, *MRPL10*, *TMEM97* y *PCTP* están localizados en el SSC12 y en muchos casos, situados próximos o incluidos en la en el intervalo de confianza del QTL. Además, todos tienen en común que se observaron diferencias de expresión en al menos uno de los análisis realizados. Asimismo, *TMEM97* y *PCTP* están relacionados con el metabolismo lipídico por lo que podrían ser considerados genes candidatos tanto funcionales como posicionales para explicar el QTL.

Es recomendable la validación mediante PCR cuantitativa de los resultados obtenidos con *arrays* de expresión ya que éstos proporcionan una medida algo grosera. Lamentablemente esta tarea no ha tenido cabida durante los años de realización de la tesis porque estos resultados se han obtenido al final de la misma, pero dicha validación será realizada en un futuro inmediato. Según lo expuesto en el párrafo anterior, la validación debería ser realizada en los genes *TMEM97* y *PCTP* para dar más solidez a su condición de genes candidatos.

Tabla 7.5. Genes diferencialmente expresados y ratios de expresión en la comparación entre genotipos del *SNP544* (*ASGA0054041A>G*)

Sonda	Anotación	AA vs GG	AA vs AG	AG vs GG
Ssc.22988.1.S1_at	<i>PCTP</i>	0.49	0.71	0.69
Ssc.5902.1.A1_at	<i>COX11</i>	1.92	1.30	1.49
Ssc.24800.2.A1_at	<i>LOC100514982</i>	1.33	1.18	-
Ssc.7209.1.A1_at	<i>COX11</i>	0.78	0.88	-
Ssc.10391.1.A1_at	<i>MCAM</i>	1.48	-	-
Ssc.8295.1.A1_at	<i>AGXT2L1</i>	0.42	-	0.46
Ssc.30718.1.A1_at	<i>KCNH2</i>	0.54	-	-
Ssc.22197.1.S1_at	<i>FANCG</i>	1.63	-	-
Ssc.15352.1.S1_at	<i>FAM117A</i>	0.70	-	-
Ssc.4110.1.S1_at	<i>TTLL4</i>	0.73	-	-
Ssc.17447.1.A1_at	<i>GTF3C5</i>	0.80	-	-
Ssc.1023.1.S1_at	<i>RBM35B</i>	0.67	-	-
Ssc.9577.1.S1_at	<i>ATXN7</i>	0.76	-	-
Ssc.2230.1.A1_at	<i>NOC3L</i>	1.34	-	-
Ssc.21245.1.S1_at	<i>GRAMD4</i>	0.78	-	-
Ssc.26209.1.S1_at	<i>SPATS2</i>	0.70	-	-
Ssc.19050.1.S1_at	<i>DHRS7C</i>	0.69	-	-
Ssc.6484.1.S1_at	<i>PNPO</i>	0.62	0.70	-
Ssc.5620.1.A1_at	<i>TOB1</i>	0.75	-	-
Ssc.21383.1.A1_at	<i>PHOSPHO1</i>	-	0.78	-
Ssc.19381.1.A1_at	<i>HOMER3</i>	-	0.83	-
Ssc.25859.1.S1_at	<i>ANKRD40</i>	-	0.79	-
Ssc.29510.1.A1_at	<i>ZNF780A</i>	-	0.76	-

DISCUSIÓN

8.1. GENES CANDIDATOS

En el presente trabajo se caracterizó el ADNc de tres genes localizados en el SSC12 en animales pertenecientes a las razas Ibérico y Landrace y se buscaron polimorfismos localizados bien en la región codificante o en regiones reguladoras de estos genes que pudiesen ser utilizados en la detección de QTLs para la composición de ácidos grasos medidos en grasa subcutánea e intramuscular así como en análisis de asociación para testar el efecto de dichos polimorfismos.

Se han seleccionado genes que codificasen para una proteína con una función relevante en el metabolismo lipídico o, en concreto con el metabolismo de ácidos grasos. Además, tomando como referencia el estudio del que parte el presente trabajo (Muñoz *et al.*, 2007), dos de los genes (*ACOX* y *ACLY*) fueron elegidos por encontrarse próximos a uno de los dos QTLs detectados.

Igualmente, a la vista de los resultados del citado estudio, se buscaron polimorfismos en las posibles regiones promotoras del gen *ACACA*, que pudiesen explicar los efectos del segundo QTL.

En mamíferos, la degradación de los lípidos se produce gracias a la acción combinada de la β -oxidación que tiene lugar en mitocondrias y peroxisomas. La función de la β -oxidación peroxisomal consiste, principalmente, en la degradación de ácidos grasos saturados de cadena larga para convertirlos en otros de cadena más corta que son exportados a las mitocondrias donde terminan el proceso de conversión a acetyl CoA (Kunau *et al.*, 1995). El primer paso de la ruta es catalizado por la enzima *ACOX1* que pertenece a la superfamilia de las flavoenzimas, a la cual también pertenecen las deshidrogenasas mitocondriales (Fournier *et al.*, 1994). *ACOX1* es la enzima limitante de la ruta y tiene por tanto un papel muy importante en el funcionamiento del proceso. De hecho, en humanos hay descrito un desorden, *Adrenoleucodistrofia pseudonatal* (Wanders *et al.*, 2001), en el cual la ausencia de la actividad de esta proteína impide la degradación de los ácidos grasos de cadena larga que se acumulan en el plasma. Este proceso desencadena una desmielinización progresiva de las neuronas que da lugar a un retraso en el desarrollo del sistema psicomotor y regresión neuronal.

El cDNA del gen *ACOX1* ha sido caracterizado en numerosas especies de mamíferos como koala, rata, ratón o humano (Fournier *et al.*, 1994; Miyazawa *et al.*, 1987; Ngo *et al.*, 2003; Nohammer *et al.*, 2000) así como en organismos modelos como *Dario rerio* o *Saccharomyces cerevisiae* (Dmochowska *et al.*, 1990; Morais *et al.*, 2007). En general, la

estructura del gen está muy conservada en las especies de mamíferos estudiadas, consta de 14 exones y 13 intrones con posición idéntica excepto en el caso del exón 14 en rata (Osumi *et al.*, 1987). Asimismo, en mamíferos y en *Danio rerio* se han descrito dos transcritos del gen *ACOX1*, que tienen la misma longitud, pero difieren en la secuencia de nucleótidos del exón 3 (3-I y 3-II), la presencia de un tipo u otro de exón se debe a corte-empalme alternativo (Fournier *et al.*, 1994; Miyazawa *et al.*, 1987; Morais *et al.*, 2007; Ngo *et al.*, 2003; Nohammer *et al.*, 2000). Estas características comunes del gen en diferentes especies son un indicador del elevado grado de conservación del mismo a lo largo de la evolución y de la relevancia de la proteína para la funcionalidad del organismo.

En el presente trabajo se caracterizó la secuencia del ADNc del gen *ACOX1* porcino. La secuencia obtenida fue alineada con las secuencias homólogas descritas en humano y ratón verificándose la presencia en porcino de los 14 exones descritos para otras especies de mamíferos. Asimismo, mediante este alineamiento se confirmó que las secuencias consenso de los extremos de los exones eran consistentes con la secuencia consenso (GT/AG). Por otro lado, también se confirmó la presencia, en hígado de las dos variantes de transcrito descritas en las especies anteriormente mencionadas.

El ADNc de *ACOX1* codifica una proteína compuesta por 661 aminoácidos, igual que ocurre en rata y en ratón y con un aminoácido menos que la proteína de humano que se compone por 660 aminoácidos. Asimismo, la secuencia aminoacídica contiene las regiones características de la proteína: la secuencia “KWWPGG” corresponde al dominio “*fatty acyl-coA oxidase*”, “CGGHGY” al motivo de unión FAD y por último, “SKL” que conforma la señal peroxisomal.

En el presente estudio se detectaron cinco polimorfismos y tres de ellos fueron genotipados en los individuos parentales. La cosegregación de los polimorfismos ha permitido definir dos haplotipos: H1: c.[481G; 1238A; *21G] y H2: c.[481C; 1238G; *21A]. Después de genotipar todos los individuos del cruce IBMAP el haplotipo H1 ha sido el más frecuente (Tabla 4.4) y no se observó ningún individuo homocigoto H2H2.

En el análisis realizado en grasa subcutánea, se detectó un QTL próximo a la posición donde mapea en el gen (17.5 cM), por lo que se realizó un análisis de asociación de estos haplotipos con los ácidos grasos afectados por el QTL (C16:0, C18:3, C20:1(n-9) y ACL). Sin embargo, no se detectó ningún efecto significativo de los haplotipos sobre estos caracteres.

A pesar de que la función de la proteína ACOX está directamente relacionada con el metabolismo de ácidos grasos y estudios realizados en ratón demuestran que la expresión

de este gen aumenta significativamente en machos alimentados con dietas ricas en grasas (Fan *et al.*, 1996), apenas se han realizado estudios de asociación de este gen con caracteres productivos en animales de granja. El único estudio publicado fue el realizado por Zuo *et al.* (2007), en el cual detectaron un efecto significativo dominante sobre el índice de firmeza y el color de la carne de un SNP localizado en el intrón 9 del gen *ACOX1*, en un cruce Large White x Meishan. Por otro lado no detectaron efectos sobre el porcentaje de grasa intramuscular.

La escasez de estudios publicados en este campo junto con los resultados obtenidos en el presente trabajo es reflejo de la dificultad de encontrar polimorfismos en este gen. En esta tesis, de 667 animales genotipados de las generaciones F2, F3, RC1 y RC2 se detectaron 560 H1H1, 107 H1H2 y ningún homocigoto H2H2. Dadas las frecuencias divergentes que, para este polimorfismo, presentan las líneas parentales, pudiera ocurrir que los individuos H2H2 posean alguna desventaja selectiva.

La enzima ATP citrato liasa (ACLY) cataliza en el citosol la formación de acetil-coenzima A y oxalacetato a partir de citrato y CoA mediante hidrólisis de ATP (Knowles *et al.*, 1974). El acetil-CoA producido está implicado en la ruta de síntesis de ácidos grasos y colesterol por lo que la ACLY es considerada una enzima lipogénica junto a la sintasa de ácidos grasos (FASN) y a la enzima acetil-CoA carboxilasa α (ACACA) (Gibson *et al.*, 1972). En mamíferos, la actividad de la enzima está regulada por la alimentación así como por los niveles de insulina (Katsurada *et al.*, 1986).

La secuencia del ADNc del gen *ACLY* ha sido objeto de numerosos estudios especialmente en rata; en esta especie el gen completo abarca 55kb organizados en 29 exones. Se han detectado dos isoformas del gen, en diferentes tejidos (cerebro, riñón, glándula mamaria, pulmón e hígado), una que contiene los 29 exones y otra que carece del exón 14 (Moon *et al.*, 1996). Otros estudios han permitido identificar también las dos isoformas con una estructura similar en humano (Suzuki y Okuda, 1981). La homología entre la secuencia de aminoácidos deducida de la proteína ACLY en humano (con 1105 aminoácidos) y en rata (con 1100 aminoácidos) es de un 96,3% (Elshourbagy *et al.* (1992)). En el presente trabajo se caracterizó la secuencia del ADNc del gen porcino *ACLY* en hígado y se observó la presencia del transcrito completo con los 29 exones, así como de la variante que carece del exón 14. Por otra parte, la longitud de la secuencia de aminoácidos deducida resultó ser ligeramente inferior a las anteriores, abarcando un total de 1076 aminoácidos.

Se detectaron y genotiparon dos polimorfismos en los individuos del cruce IBMAP pero, tan sólo se detectó una asociación sugestiva entre el polimorfismo *ACLY*: c.523T>C y el porcentaje del ácido palmítico (C16:0) medido en grasa subcutánea, con un valor de 0.182 ± 0.099 ($P < 0.059$). Hasta la fecha, no se ha publicado en porcino ningún estudio en el que se analice la implicación de este gen en la variación del perfil de ácidos grasos o algún otro tipo de carácter relacionado la calidad de carne. Sin embargo, se han publicado numerosos trabajos en los que se describe una regulación de este gen a nivel transcripcional en diferentes especies y se relacionan dichos niveles transcripcionales con un mayor o menor contenido en grasa del organismo estudiado. Estudios realizados en gallinas describen la presencia de niveles significativamente superiores de ARNm en hígado de líneas de gallinas grasas respecto a otras líneas magras (Daval *et al.*, 2000). Asimismo Ren *et al.* (2008) describieron una mayor expresión del gen en cerdos de la generación F1 de cruces Meishan x Large-White y Large-White x Meishan, respecto a los individuos parentales.

Uno de los elementos reguladores de la expresión de este gen descritos hasta el momento es la proteína codificada por el gen *Sterol Regulatory Element Binding Factor 1* (*SREBF1*). Esta proteína actúa como factor de transcripción regulando la expresión de proteínas lipogénicas como *ACLY* y *ACACA* (Sato *et al.*, 2000). Asimismo, la actividad de la propia proteína *SREBF1* es inducida mediante la insulina, de tal manera que cuando el organismo ingiere alimentos, se produce un incremento de secreción de insulina y se desencadena una cascada que deriva en un aumento de la síntesis *de novo* de ácidos grasos y colesterol (Foretz *et al.*, 1999).

En porcino, el gen *SREBF1* mapea en el extremo del brazo largo del cromosoma 12, pero no está localizado dentro de ninguno de los intervalos de confianza de los QTLs detectados (Muñoz *et al.*, 2007). Por lo tanto, la elección de este gen candidato fue realizada en base a sus propiedades funcionales y a la regulación transcripcional de los genes que sí están localizados en los intervalos de confianza y que codifican las proteínas lipogénicas *ACLY* y *ACACA* mencionadas anteriormente. Asimismo, su posición en el cromosoma, ha permitido disponer de un marcador en una región que la que en el análisis inicial no se pudo disponer de ningún microsatélite informativo.

En el presente estudio se detectaron siete polimorfismos localizados en el ADNc del gen *SREBF1* y, de ellos, los SNPs c.1023T>C y c.2597C>T fueron los más informativos. El polimorfismo c.2597C>T da lugar a un cambio de aminoácido que, según los modelos generados por las herramientas SMART y PSIPRED, estaría localizado en una

región de baja complejidad. Asimismo este cambio sería tolerado por la estructura de la proteína, en base a las predicciones de la herramienta SIFT. En lo que concierne al SNP c.1023T>C las predicciones del *RNA Secondary Structure Prediction* indican que podría inducir cambios estructurales en la conformación secundaria del RNA. Chen *et al.* (2008) analizaron este polimorfismo en las razas puras de cerdos Eurhalian y Sutai y detectaron un efecto sobre el porcentaje de grasa intramuscular de una de las clases de los homocigotos respecto al resto de genotipos.

En nuestro trabajo se ha incluido el SNP c.1023T>C como marcador en la detección de QTLs en grasa subcutánea y en una de las detecciones llevadas a cabo en grasa intramuscular. Asimismo, se llevó a cabo el análisis de asociación del SNP sobre los diferentes porcentajes de ácidos grasos de ambos tejidos (resultados no mostrados), aunque no se detectó asociación significativa con ninguno de los caracteres. Por otra parte un trabajo realizado en ganado vacuno japonés, describe una inserción/delección localizada en un intrón de este gen con efectos sobre la proporción de MUFA en grasa intramuscular (Hoashi *et al.*, 2007). Dichos efectos han sido contrastados en diferentes líneas genéticas de bovino, obteniéndose resultados dispares. Por un lado se detectó un efecto en la concentración de los ácidos esteárico y linoleico así como en el porcentaje de ácidos polinsaturados medidos en la grasa intramuscular de la carne de vacuno coreano (Bhuiyan *et al.*, 2009). Sin embargo en otras poblaciones de ganado vacuno japonés así como en toros Fleckvieh, no se detectaron efectos relevantes en el perfil de ácidos grasos de las grasas subcutánea e intramuscular (Barton *et al.*, 2010; Ohsaki *et al.*, 2009).

Nuestros resultados no proporcionan indicios de que pueda haber otro polimorfismo en el gen que estuviese asociado al perfil de ácidos grasos de ambos tejidos analizados, no obstante la relevancia funcional del gen, su relación con las enzimas lipogénicas anteriormente citadas y los resultados publicados en bovino, dejarían la puerta abierta a una continuación del análisis de este gen. De hecho, se podría comprobar si alguno de los otros polimorfismos que hemos detectado tiene algún efecto sobre la composición de ácidos grasos.

La enzima acetil-CoA carboxilasa α cataliza la formación de malonil-CoA a acetil-CoA en el citoplasma, el paso limitante en la síntesis de los ácidos grasos de cadena larga. En humanos el gen *ACACA* está compuesto por 58 exones principales y, de ellos, los exones E5 a E58 codifican para una proteína de 265 KDa. En un estudio previo realizado en nuestro grupo (Muñoz *et al.*, 2007) el ADNc de los exones E5 a E58 del gen *ACACA*

porcino fue caracterizado detectándose 15 polimorfismos sinónimos y efectos significativos de dos de ellos sobre la composición en ácidos grasos de la grasa dorsal.

En este trabajo se ha ampliado la búsqueda de polimorfismos a las posibles regiones reguladoras del gen. Aunque no se dispone de información referente a porcino, se sabe que en otras especies de mamíferos la transcripción del gen *ACACA* está controlada por tres promotores (PI, PII y PIII) y da lugar a una elevada heterogeneidad en la región 5' no traducida (Mao et al. 2001). En bovino, los transcritos controlados por PI se expresan principalmente en tejido adiposo pero también en hígado en condiciones lipogénicas. PII es una región rica en dinucleótidos CpG que parece estar funcionalmente conservada en todas las especies de mamíferos. Los islotes CpG son característicos de los genes que se expresan de manera constitutiva y todo indica que PII actúa como un “*Promotor Housekeeping*” aunque puede ser estimulado en glándula mamaria, durante la lactación. El tercer promotor, PIII es activo en glándula mamaria. En humanos también se conocen tres promotores (Mao et al., 2003; Barber et al., 2004) pero, en este caso, PI no presenta homología con las secuencias de PI descritas en rata, bovino u ovino y parece actuar como un promotor constitutivo. Por otra parte la actividad de PIII aumenta en líneas celulares de tumores mamarios y se cree que, en humanos tal como en rumiantes, podría desempeñar una función importante en la producción de grasa de la leche, necesaria para el desarrollo neonatal. Los transcritos derivados de PI y PII dan lugar a proteínas idénticas (265 kDa) ya que al codón ATG de inicio de la traducción se encuentra en el exón 5. PIII transcribe el ARNm desde el exón 5A, localizado en el intrón 5 (véase figura 5.4), y da lugar a una proteína de 259 kDa. La secuencia comprendida entre los exones 1 y 5A abarca 110 kb, un tercio de la longitud total del gen *ACACA* y está considerada como la región reguladora del gen.

Dada la dimensión de este gen y las características de los promotores, este trabajo se ha centrado en la búsqueda de polimorfismos en la región correspondiente a PI. Asimismo, un trabajo reciente, desarrollado por Zhang *et al.* (2010) en vacuno, describe asociaciones significativas entre polimorfismos identificados en el promotor I y diferentes ácidos grasos en el músculo *longissimus dorsi*. Se utilizaron las secuencias correspondientes a PI identificadas en humano y en bovino para buscar secuencias homólogas en porcino, depositadas en las bases de datos. El hecho de que, en la búsqueda realizada en las bases de datos, no hayamos encontrado secuencias con homología al PI de humano sino al de bovino concuerda con lo que se ha dicho anteriormente. Sorprendentemente la región identificada en *Sus scrofa* se encuentra entre los exones “1” y “1c-like” cuando cabría esperar que la región promotora estuviera por encima del primer exón, aunque de acuerdo con

Mao *et al.* (2003), está dentro de la gran región reguladora del gen. Se ha analizado el polimorfismo *ACACA*: g.48001A>G y se ha detectado una asociación significativa con el ácido esteárico (C18:0) de la grasa dorsal pero no se han detectado asociaciones significativas para ningún ácido graso de la grasa intramuscular.

El gen *ACACA* abarca un tamaño de más de 214000 pb en humano y, en porcino quedan muchas regiones pendientes de analizar. No obstante las nuevas técnicas de secuenciación masiva así como un conocimiento más preciso del genoma porcino (se prevé que la nueva versión Ssc10 esté disponible en la base de datos Ensemble en primavera del presente año), facilitan la tarea de detectar nuevos polimorfismos en este gen.

8.2. DETECCIÓN CLÁSICA DE QTLs PARA COMPOSICIÓN DE ÁCIDOS GRASOS EN GRASA SUBUCTÁNEA

El nuevo análisis de detección de QTLs para ácidos grasos en grasa subcutánea llevado a cabo en el presente estudio contó con los genotipos de tres nuevos marcadores (SNPs *ACLY*: c.*523 T>C y *SREBF1*: c.1023 T>C y haplotipos H1 y H2 detectados en el gen *ACOX1*: c.[481G>C; 1238A>G; *21G>A]). Además de intentar reducir los intervalos de confianza de los QTLs detectados en el trabajo previo (Muñoz *et al.*, 2007), se quiso analizar si los polimorfismos detectados en los genes *ACOX1* y *ACLY* explicaban los efectos del QTL1.

En primer lugar, al construir el mapa de ligamiento con la información de estos tres nuevos marcadores, la longitud del mapa de ligamiento del SSC12 porcino pasó de medir 117 cM a 136.2 cM, disminuyendo la distancia entre marcadores de 12.4 cM a 10.48 cM. Al igual que en el trabajo previo se detectaron dos QTLs, el primero de ellos (QTL1) delimitado por los marcadores *ACOX1* y *ACLY*. La amplitud de este QTL es efectivamente inferior al detectado por Muñoz *et al.* (2007) ya que éste estaba acotado por los marcadores *s0143* (posicionado antes que el gen *ACOX1* en el nuevo mapa de ligamiento) y más allá del marcador *sur1307* (posicionado después del gen *ACLY* en el nuevo mapa de ligamiento). El nuevo QTL1 tiene efecto sobre los porcentajes de ácido linolénico (C18:3, n-3), gadoleico (C20:1, n-9) y palmítico (C16:0) así como sobre la Longitud Media de la Cadena (ACL).

Por otra parte, el QTL2, está delimitado por los marcadores *sur1802* y *ACACA* y al igual que en el caso anterior, el intervalo de confianza es inferior al detectado por Muñoz *et al.* (2007). Asimismo, se detectaron efectos del QTL sobre los porcentajes de los ácidos

grasos mirístico (C14:0), palmitoleico [C16:1, n-9], esteárico (C18:0) y vaccénico [C18:1, n-7], siendo el efecto sobre C14:0 principalmente dominante. Los efectos detectados de los dos QTLs fueron comunes en ambos trabajos en todos los casos excepto en el caso del C14:0 ya que Muñoz *et al.* (2007) detectaban un efecto sobre el mismo principalmente aditivo y localizado en el QTL1 y en el presente estudio se detectó un efecto principalmente dominante y localizado en el QTL2. Por otro lado se detectó un nuevo efecto dominante del QTL1 sobre el C18:3 (n-3) no detectado en el trabajo previo.

Los análisis de asociación con los genes candidatos pusieron de manifiesto, como se ha comentado previamente en la discusión de la caracterización de genes candidatos, que los polimorfismos detectados en los genes *ACOX1* y *ACLY* no explican los efectos observados del QTL1, por lo tanto, un nuevo gen candidato localizado dentro del intervalo de confianza del QTL debe ser buscado para explicar los efectos del mismo.

Por otro lado, se repitieron los diferentes análisis de asociación (modelo animal y modelo de asociación asistida por marcadores) entre el polimorfismo *ACACA*: c.5634T>C y los porcentajes de los ácidos grasos palmitoleico [C16:1, n-9], esteárico (C18:0) y vaccénico [C18:1, n-7] y se observó que los efectos de este polimorfismo se mantienen. Asimismo, se realizó un nuevo análisis de asociación entre el nuevo SNP intrónico detectado en este gen (g.48001 A>G) observándose un efecto del mismo sobre el porcentaje de C18:0. A pesar de estos resultados significativos, no se puede afirmar que alguno de estos polimorfismos sea el responsable del efecto detectado. Se realizaron experimentos y estudios complementarios que proporcionasen evidencias sobre la relevancia de estos polimorfismos.

En el primero de estos estudios, que hubo de realizarse en muestras de animales del RC2, se analizó mediante PCR cuantitativa el efecto de los genotipos de estos SNPs sobre la expresión del gen *ACACA* en grasa subcutánea. Para ambos polimorfismos se encontraron diferencias de expresión entre los genotipos, significativas para el SNP g.48001 A>G y sugestivas para el c.5634T>C. Se aprecia un patrón de dominancia en ambos casos ya que la expresión de uno de los homocigotos difiere significativamente de los demás genotipos. Esta dominancia del efecto sobre la expresión no es concordante con la naturaleza aditiva del efecto de estos polimorfismos sobre la composición de ácidos grasos.

Paralelamente, se llevó a cabo un segundo estudio para comprobar si los efectos del SNP c.5634T>C se mantenían en dos poblaciones no experimentales *Ibérico x Ibérico* y *Duroc x Ibérico*. En la primera de ellas, se detectaron efectos significativos del SNP sobre los porcentajes de los ácidos grasos mirístico (C14:0) y palmítico (C16:0), pero no sobre

C16:1(n-9), C18:0 y C18:1(n-7). Sin embargo, en los animales cruzados *Duroc* x *Ibérico* no se detectaron efectos significativos del SNP sobre ninguno de los ácidos grasos analizados.

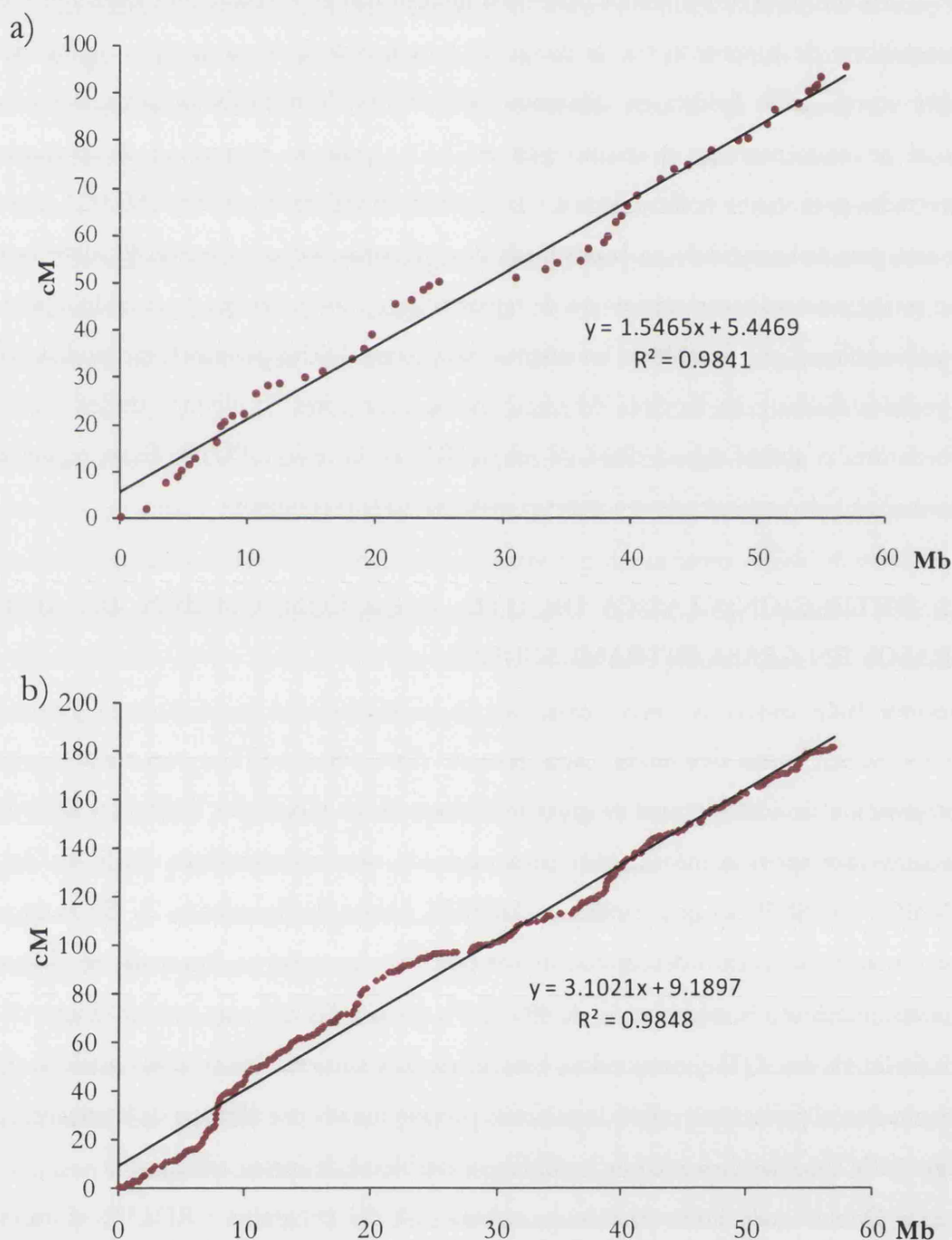
El conjunto de resultados obtenidos indican que el polimorfismo analizado está en desequilibrio de ligamiento con la mutación causal, dada la inconsistencia de los efectos sobre composición lipídica en diferentes poblaciones. Muy probablemente la mutación causal se encuentre en el mismo gen *ACACA*, dada la reiteración de asociaciones detectadas para varios polimorfismos y la función crucial de la proteína *ACACA*. Aunque en esta tesis se ha incidido en la búsqueda de polimorfismos en la región PI, debe notarse que está pendiente la caracterización de regiones importantes del gen que podrían albergar el polimorfismo responsable de los efectos detectados. Dicho polimorfismo podría afectar a posibles dianas para factores de transcripción/represores en algunos de los otros dos promotores o alterar alguna diana de microARN en la región 3'UTR. Estas regiones del gen no han sido secuenciadas en nuestro material hasta el momento.

8.3. DETECCIÓN CLÁSICA DE QTLs PARA COMPOSICIÓN DE ÁCIDOS GRASOS EN GRASA INTRAMUSCULAR

Se realizaron dos análisis sucesivos de detección de QTLs con efectos sobre la composición de ácidos grasos de grasa intramuscular en individuos F3, RC1 y RC2. En el primero se empleó la información combinada de siete microsátélites, 26 SNPs del *chip SNP60* y un SNP del gen candidato *SREBF1* (mapa de ligamiento 2). En la segunda aproximación se intentó definir con mayor precisión las regiones detectadas en el anterior análisis, utilizando la información de 489 SNPs del *chip SNP60* para reducir el intervalo de confianza de los QTLs detectados. Una diferencia adicional entre estos análisis es que mientras en el primero se utilizó la posición proporcionada por el mapa de ligamiento, en el segundo la localización en cM de los SNPs se estableció de forma aproximada.

Debido a la limitación de la versión 2.4 del programa CRIMAP, el mapa de ligamiento 3 se obtuvo con una información reducida de 50 SNPs seleccionados entre los 489. Para la detección de QTLs, estos marcadores se posicionaron uniformemente a lo largo del SSC12 a una distancia de 0.2 cM, que corresponde a la distancia media estimada a partir de la relación entre el mapa 3 y el mapa físico (véase Resultados 6.2.2).

Figura 8.1. Regresión lineal de la posición genética (cM) sobre la posición física (Mb) calculada a partir de los mapas obtenidos con a) 50 SNPs y b) 489 SNPs



Afortunadamente, al finalizar este trabajo encontramos disponible en la red la versión 5.2 del programa CRIMAP, en la cual sí es posible realizar un mapa genético con una densidad superior de marcadores. Ello nos ha permitido verificar el alcance de esta aproximación sobre los resultados obtenidos tanto en el mapeo de ligamiento como en la detección de QTLs. Para ello se hizo un nuevo mapa, con los 489 SNPs utilizados en la detección de QTLs (mapa 4). Este mapa de ligamiento fue construido posicionando inicialmente los marcadores en el mismo orden que en el mapa físico. La longitud resultante fue de 181.40

cM y la distancia media entre marcadores de 0.37 cM. Este valor indica que la distancia entre marcadores estimada del mapa 3 (0.2 cM) estaba subestimada.

Mediante el cálculo de la regresión lineal entre la posición en cM de cada SNP en el mapa de ligamiento del SSC12 y su localización en Mb en el mapa físico se realizó una estima de la relación entre las distancias genéticas y número de Megabases. Este cálculo se llevó a cabo con la información de los mapas de ligamiento 3 y 4. En ambos casos, la regresión lineal presenta un buen ajuste ($R^2 = 0.98$), sin embargo, en el mapa 4 la equivalencia Mb-cM duplica la del mapa 3. De cualquier forma estas regresiones se alejan de la convención aceptada $1\text{cM} = 1\text{ Mb}$ (Hayes *et al.*, 2003; Tenesa *et al.*, 2007), con tasas de recombinación hasta tres veces superiores a las esperadas. Por otra parte, el ajuste de ambas rectas es mejor desde la posición de 40 Mb hasta el final del cromosoma y ambas también presentan un peor ajuste en la región central del mismo, con una pendiente algo inferior. Este resultado podría reflejar la menor tasa de recombinación alrededor de la región centromérica.

Como la distancia media entre marcadores del nuevo mapa es superior a la anteriormente calculada, se realizó una nueva detección de QTLs con los mismos 489 marcadores pero considerándolos uniformemente separados por 0.37 cM. El mayor tamaño en cM del cromosoma supuso la realización de un mayor número de tests de presencia/ausencia del QTL (1 test/cM): 181 en lugar de los 98 realizados en el anterior análisis. Los valores de FDR calculados de acuerdo con Benjamini y Yekutieli (2005) fueron por ello más restrictivos, de tal manera que para $\text{FDR} = 10\%$ (QTL sugestivo) el umbral de P corresponde a $0.1/181 = 5.52 \times 10^{-4}$ y para $\text{FDR} = 5\%$ (QTL significativo), el valor de P corresponde a $0.05/181 = 2.76 \times 10^{-4}$.

En la tabla 8.1 se comparan los resultados del nuevo análisis de detección de QTLs obtenidos a partir de las distancias media entre SNPs del mapa 4 con los resultados significativos obtenidos en el análisis realizado con el mapa 3. Como consecuencia del mayor tamaño del mapa 4, se modifican todas las posiciones expresadas en cM. La posición del QTL y su efecto sobre el porcentaje de C16:0 son similares en ambos análisis. Los efectos sobre los índices UI y SFA y su significación nominal son algo inferiores en el análisis realizado con la distancia estimada del mapa 4. La posición correspondiente al máximo valor de LR para el ácido graso C20:4 y los índices DBI y PUFA/SFA cambia de 31 a 19 Mb en el análisis más reciente con la distancia media entre marcadores de 0.37 cM.

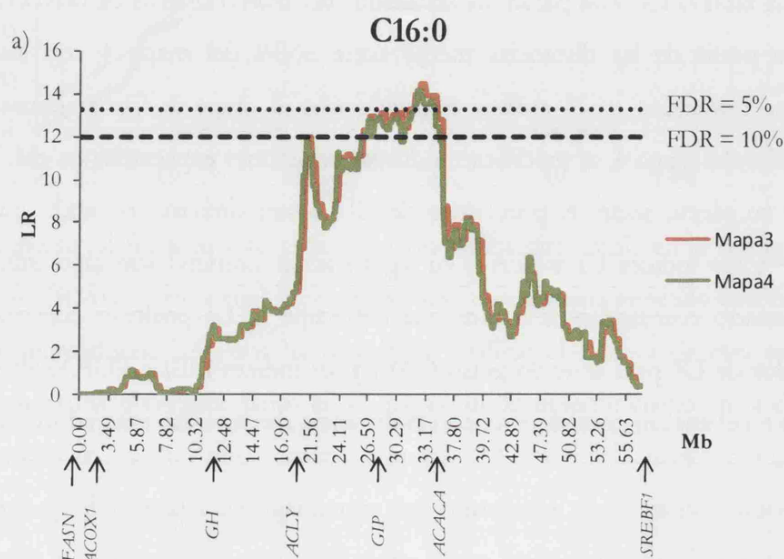
DISCUSIÓN

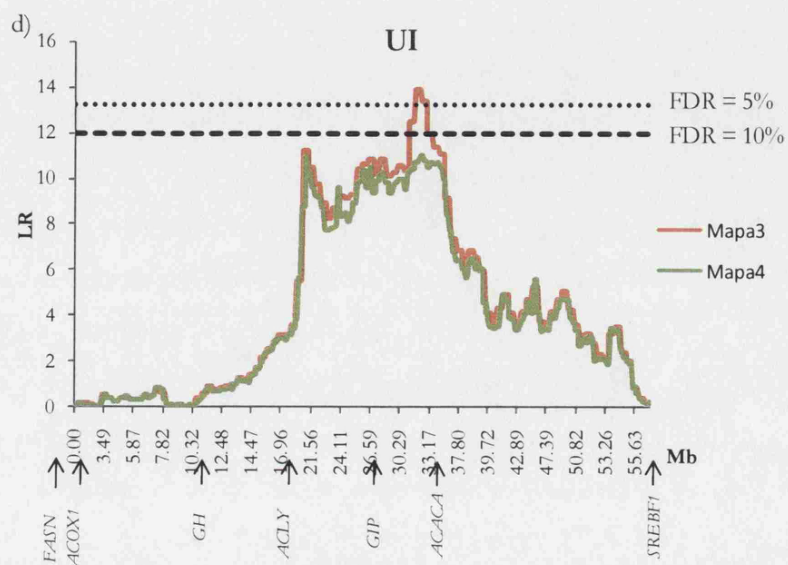
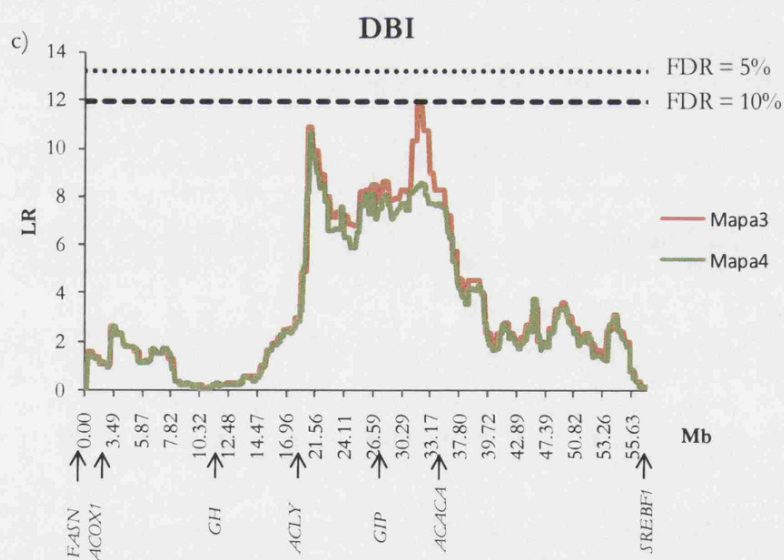
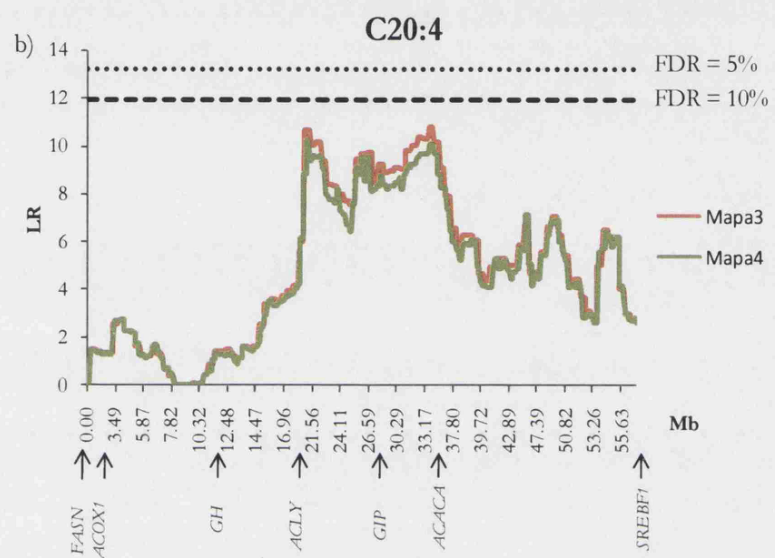
Tabla 8.1. Comparación entre los resultados significativos y sugestivos obtenidos en la detección de QTLs, utilizando los mapas 3 y 4 con respectivas distancias entre los 489 SNPs de 0.20 cM y 0.37 cM.

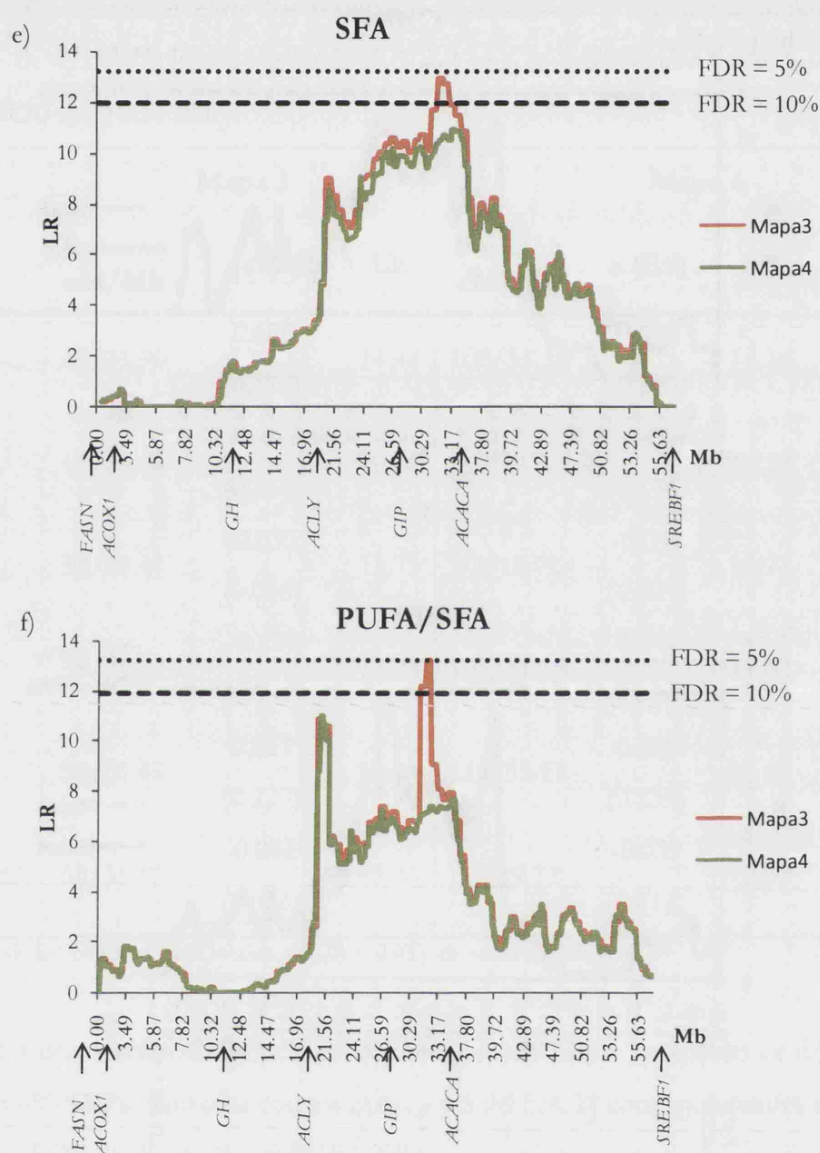
Carácter	Mapa 3			Mapa 4			P
	Posición cM/Mb	a (ES)	LR	Posición cM/Mb	a (ES)	LR	
C16:0	59/31.90	0.667 (0.174)	14.48	108/31.55	0.683 (0.179)	14.16	1.6 x10 ⁻⁴
C20:4	61/34.08	-0.443 (0.134)	10.86	72/19.72	-0.438 (0.136)	10.28	0.0013
DBI	58/31.47	-0.033 (0.008)	11.73	72/19.72	-0.031 (0.009)	10.71	0.0017
UI	58/31.47	-0.129 (0.034)	13.93	108/31.55	-0.120 (0.036)	11.02	9.0 x10 ⁻⁴
SFA	58/31.47	0.827 (0.227)	12.94	111/33.17	0.800 (0.239)	10.97	9.2 x10 ⁻⁴
PUFA/SFA	58/31.47	-0.042 (0.014)	13.26	72/19.72	-0.039 (0.014)	10.00	9.1 x10 ⁻⁴

Los valores de LR en rojo corresponden a FDR < 0.05 y en verde a FDR < 0.10

Figura 8.2. Comparación de los perfiles de los valores de LR de los análisis de detección de QTLs con 489 SNPs realizados con los mapa 3 y 5 del SSC12 correspondientes a: a) C16:0, b) C20:4, c) DBI, d) UI, e) SFA, f) PUFA/SFA.







En la figura 8.2 se representan los diferentes perfiles de verosimilitud de los análisis de detección de QTLs realizado con los mapas 3 y 4. Como los mapas de ligamiento tienen longitudes diferentes, se ha representado la variación del valor de LR a lo largo de la longitud del SSC12 expresado en Mb. Los perfiles correspondientes a los ácidos grasos C16:0 y C20:4 son similares (Figuras 8.2. a y b) aunque, como se ha comentado en el caso del ácido C20:4, se ha modificado la posición del posible QTL. Sin embargo, los perfiles de verosimilitud correspondientes a los diferentes índices (Figuras 8.2 c, d, e y f) presentan diferencias más notables entre ambos análisis. Varios de estos perfiles sugieren la posible presencia de dos QTLs, aunque esta hipótesis fue rechazada en el análisis de comparación de modelos de uno y dos QTLs (Tabla 6.7).

En los últimos años, se ha considerado que uno de los inconvenientes de realizar análisis de detección de QTLs es debido a que las regiones con efecto sobre los caracteres suelen ser extensas y es complicado conseguir intervalos de confianza pequeños que acoten los QTLs detectados y faciliten el proceso de búsqueda de la mutación causal (Fan *et al.*, 2010). En el presente estudio se pretendía mejorar la precisión respecto al análisis realizado con información combinada introduciendo una mayor cantidad de marcadores. El QTL con efecto sobre C16:0 tiene un intervalo de confianza de 45 a 52 cM en el análisis con el mapa 2 (información combinada), mientras que en el realizado a partir de la información con los mapas 3 y 4, los intervalos de confianza estimados abarcan desde la posición 26.47 a 35.11 Mb y de 30.96 a 35.02 Mb respectivamente. Aunque no es sencilla la comparación entre estos intervalos, no parece que se haya obtenido una mejora notable en la precisión de la localización de este QTL. Posiblemente el número de recombinantes presentes en este material biológico sea un aspecto más crítico que el número de SNPs, especialmente cuando el análisis realizado con el mapa 2 utilizaba ya un importante número de marcadores genéticos (27 SNPs bialélicos y seis microsatélites multialélicos).

En este mismo material biológico, Fernández *et al.* (2011) realizaron un análisis clásico de detección de QTLs para diversos caracteres en los cromosomas 1, 10 y 11. Sus resultados han permitido la detección de nuevos QTLs, que no se habían detectado en análisis anteriores, realizados con mapas de baja densidad basados en microsatélites. Esos resultados han sido verificados en estudios paralelos de asociación genómica en los mismos cromosomas (Ramayo-Caldas *et al.*, 2010). Este conjunto de resultados parece indicar que la detección de QTLs basada en información de un gran número de SNPs puede ser relevante si hace posible el mapeo de regiones del genoma pobremente cubiertas por otros marcadores. En casos como el del SSC12 porcino, en que se disponía de una buena cobertura previa con microsatélites y SNPs, la información añadida no conlleva la detección de nuevos QTLs.

Finalmente, cabe destacar que en los tres análisis se detectó un QTL con efecto sobre C16:0. El ácido graso C16:0 es uno de los mayoritarios, presentando una media de un 22.7% del total de ácidos grasos en grasa intramuscular. Los efectos detectados sobre el mismo son de una magnitud del 0.6% por copia del alelo *Q* Ibérico, equivalentes a un 2.6% de la media del carácter. Los porcentajes relativos al total de ácidos grasos son menores para los otros ácidos para los que se detectó el posible efecto de algún QTL, lo que dificulta su detección en experimentos como el nuestro de no excesiva potencia.

Ledur *et al* (2010) establecieron que para que un número elevado de marcadores incremente la potencia de un análisis de detección de QTLs es importante que los marcadores se dispongan de manera equidistante a lo largo del genoma. Además, establecen que para que estos análisis sean más efectivos, un elevado número de SNPs (superior al disponible en el *chip* de porcino) debe ser utilizado. Sin embargo, determinan como factor limitante tener un tamaño poblacional elevado, lo cual no ocurre en este estudio ya que se contaba con un tamaño poblacional de unos 200 individuos (varía en función del ácido graso analizado) lo cual, disminuye considerablemente la potencia del experimento. Asimismo la procedencia de retrocruces de la mayoría de animales analizados contribuye a disminuir la potencia de nuestro experimento en el que los animales *QQ* están escasamente representados.

Por otra parte, en el diseño del chip *SNP60* se decidió eliminar todos aquellos polimorfismos para los cuales se había descrito efecto sobre algún carácter, ya que el *chip* estaba enfocado, en un principio, a la realización de estudios de variabilidad entre diferentes razas (Ramos *et al*, 2009). Este hecho podría reducir su utilidad en el análisis de detección clásica de QTLs.

Finalmente, se construyó un último mapa a partir de la información de los 978 SNPs de este cromosoma con $MAF > 0.15$. Las frecuencias de recombinación estimadas entre los SNPs *ASGA0104321* y *MARC0110971* y sus contiguos presentaron valores muy elevados (0.28 y 0.42), que indican altamente improbable su localización en el SSC12 (Figura 8.3). Tras la eliminación de esos dos SNPs, el mapa resultante tuvo una longitud de 285.04 cM y la distancia media entre marcadores fue de 0.29 cM, lo que supone un valor intermedio entre las dos distancias anteriormente estimadas.

Figura 8.3. Fragmentos de la construcción del mapa 4 en el cual se marcan en rojo los dos SNPs descartados por sus frecuencias de recombinación elevadas respecto a su supuesta localización física

Sonda	Frac.Recomb.	cM	Sonda	Frac.Recomb.	cM
134	MARC0034050	0.000 0.000	974	MARC0114025	0.000 0.000
135	ASGA0104321	0.276 31.079	975	H3GA0035092	0.000 0.000
136	ALGA0064767	0.264 29.383	976	ASGA0055477	0.000 0.000
137	H3GA0033394	0.004 0.436	977	MARC0110971	0.418 60.322
		0.013 1.302			

El hecho de encontrar SNPs mal localizados en el mapa físico pone de manifiesto la existencia de errores en la versión utilizada del mismo (Ssc9). Estos errores no son de

gran relevancia en este caso pero, en el otro estudio realizado en este material hasta la fecha, se han detectado bloques de varios SNPs mal mapeados (Fernández *et al.*, 2011). Por tanto, hay que ser cautos a la hora de asignar una posición como verdadera dentro de un cromosoma, especialmente en los análisis de asociación en los cuales no se realizan este tipo de comprobaciones.

8.4. ANÁLISIS DE ASOCIACIÓN BASADOS EN EL *chip* SNP60

En estos últimos años se han publicado estudios de asociación a nivel genómico utilizando datos de genotipado masivo, basados en el desequilibrio de ligamiento entre marcadores y mutación causal. Estos trabajos han sido particularmente abundantes en la especie humana, pese a que el desequilibrio de ligamiento es menor en ésta que en animales domésticos, debido lógicamente al interés en el conocimiento de la base genética de las diferentes patologías (obesidad, cáncer, psoriasis etc) por lo que este tipo de plataformas de genotipado se desarrollaron inicialmente para estudios de biomedicina.

Los estudios de asociación a nivel genómico (o cromosómico) en animales domésticos plantean una dificultad derivada de que su compleja estructura de parentesco (en nuestro caso, mezcla de familias de medios hermanos y hermanos completos) genera desequilibrio de ligamiento aún entre genes no ligados. Este problema se ha intentado atenuar realizando una corrección previa de los fenotipos por un efecto familia (Aulchenko *et al.*, 2007), o mejor aún utilizando un modelos animal que tiene en cuenta el conjunto de relaciones de parentesco e incorpora un efecto poligénico que corresponde a la acción aditiva de otros genes que afectan el carácter analizado. La omisión de este término en el modelo estadístico puede hasta duplicar el número de falsas asociaciones positivas (Goddard y Hayes, 2009). Resultados de estudios de simulación confirman que los modelos que tienen en cuenta el parentesco entre los animales estudiados dan lugar a un número menor de falsas asociaciones que los modelos que lo ignoran, aunque la potencia de éstos últimos pueda ser algo superior (Erbe *et al.*, 2011).

Uno de los problemas que plantean estos análisis es la elección de un método adecuado de corrección de la multiplicidad de tests, que permita simultáneamente reducir la proporción de falsos positivos debido a la realización de múltiples tests y el riesgo de rechazo de falsos negativos. En algunos estudios (Gao *et al.*, 2010; Lantieri *et al.*, 2009), se han utilizado tests de permutación para corregir la significación de los tests, éstos son un excelente método de fijación de umbrales en poblaciones no relacionadas como es el caso

de las humanas, sin embargo, en poblaciones con relaciones de parentesco, los genotipos generados al azar en las diferentes permutaciones no preservan la estructura de genealogía que existe en los datos reales. Otros autores, optan por aplicar la corrección de Bonferroni (Lander y Kruglyak, 1995) para ajustar los niveles de significación genómica (Li *et al.*, 2010; Simon-Sanchez *et al.*, 2009; Uemoto *et al.*, 2010), esta corrección consiste en dividir el nivel de significación estadística, para un solo test de la hipótesis nula, por el número de tests realizados. Hemos utilizado una variante de esta corrección en el estudio previc de detección clásica de QTLs. No obstante, esta corrección se considera muy restrictiva en análisis de asociación, ya que el desequilibrio de ligamiento entre SNPs, hace que el número efectivo de tests independientes sea inferior al de SNPs considerados en el análisis de asociación.

En el presente estudio se ha utilizado una extensión del estadístico tasa de falso descubrimiento (FDR) (Benjamini y Hochberg, 1995) denominado *q-value* (Storey y Tibshirani, 2003), en el cual a cada valor de p le corresponde un valor de q que permite identificar como significativos los inferiores a la tasa de FDR escogida ($q < \text{FDR}$), este cálculo es realizado en base a la proporción de verdaderas hipótesis nulas a partir del histograma del conjunto de valores p obtenidos. Este tipo de test es recomendado para análisis de asociación genómica ya que el algoritmo de cálculo requiere histogramas basados en un elevado número de valores p . Esta corrección fue utilizada, entre otros, en uno de los dos estudios de asociación genómica publicados en porcino (Duijvesteijn *et al.*, 2010). En el presente estudio, realizado a nivel cromosómico, se consideraron como significativos aquellos SNPs cuyo valor de q fuese inferior a 0.10.

Este análisis de asociación a partir de los datos de genotipado del *chip SNP60* perseguía un doble objetivo. A partir de la detección de SNPs con efecto sobre los caracteres analizados, se trataba de localizar nuevos genes candidatos o bien detectar efectos de SNPs no genotipados hasta el momento en los genes candidatos propuestos anteriormente para el SSC12. Asimismo, se pretendía contrastar la detección de QTLs realizada anteriormente con los resultados obtenidos con otra aproximación metodológica.

En los diferentes análisis realizados, se detectaron efectos significativos de varios SNPs sobre el ácido graso C16:0 y el índice C20:4/C18:2. Asimismo, se detectaron efectos de otros SNPs sobre el ácido graso C20:4 y los índices PI y PUFA/SFA, aunque no llegaron a superar ($q < 0.15$) la tasa establecida de falso descubrimiento.

El análisis de asociación del contenido en ácido palmítico C16:0 es el que ha dado lugar a la detección del mayor número de SNPs (167) con efectos significativos ($q < 0.10$).

Cercanos al intervalo de confianza del QTL previamente detectado en el análisis clásico (30.96 -35.02 Mb) o incluidos dentro del mismo, fueron detectados cuatro SNPs localizados en dos genes (*ACACA* y *PCTP*) con función relacionada con el metabolismo de ácidos grasos. El SNP 636 (*ALGA0066302*), localizado en la posición 36.94 Mb correspondiente a un intrón del gen *ACACA*, tiene un efecto de 0.400 ± 0.156 % sobre C16:0, equivalente a 1.8% de la media del carácter. Gallardo *et al* (2009) realizaron en una población Duroc un estudio de asociación entre varios SNPs detectados en el gen *ACACA* y el contenidos de diferentes ácidos grasos en grasa intramuscular. Como resultado, detectaron cuatro SNPs sinónimos con efecto sobre el C16:0. Estos resultados -que corresponden a SNPs diferentes de los analizados en esta tesis- refuerzan en otro material biológico la hipótesis de que el gen *ACACA* es un potente gen candidato para albergar la mutación causal.

Los SNPs 542, 544 y 545 (*ASGA0054039*, *ASGA0054041* y *ASGA0054044*), anotados en el gen *phosphatidylcholine transfer protein* (*PCTP*) localizado en 29.74 Mb tienen efectos aditivos estimados del 0.265 ± 0.182 , 0.349 ± 0.129 y 0.398 ± 0.156 sobre el porcentaje de ácido palmítico C16:0, equivalentes al 1.9, 1.4 y 1.5% de su media, respectivamente. Estos SNPs están situados en el intervalo de confianza del QTL anteriormente detectado. El gen *PCTP* codifica para una proteína de la superfamilia START, que son proteínas reguladoras esteroideogénicas. Estas proteínas se unen a los lípidos y regulan el transporte de los mismos en los diferentes tejidos (Wirtz, 1991). Estudios realizados *in-vitro*, han demostrado que la función de la proteína PCTP consiste en el transporte intermembranal de fosfatidilcolinas, no obstante, este mecanismo no ha podido ser demostrado *in-vivo* (Kanno *et al.*, 2007). Asimismo, ratones portadores de la mutación puntual Arg120His, en este gen, tenían alterado el metabolismo de fosfatidilcolinas (Pan *et al.*, 2006). Estudios recientes han revelado que PCTP es una enzima clave que controla el metabolismo lipídico y de la glucosa y que su expresión está regulada por la proteína *PPARA*. Asimismo, limita el acceso de los ácidos grasos a la mitocondria para que participen del proceso de β -oxidación mitocondrial (Kang *et al.*, 2010a; Kang *et al.*, 2010b). Hasta el momento, no se ha publicado ningún estudio de este gen o de la proteína que codifica en porcino ni en otras especies domésticas.

Además de los efectos citados sobre el C16:0 de los polimorfismos del gen *PTCP*, se detectaron efectos sugestivos del SNP542 sobre C20:4 y del SNP544 sobre C20:4, PI y PUFA/SFA. Estos resultados indican que este gen es un sólido candidato posicional, aunque se desconoce gran parte de su función fisiológica en animales domésticos. Como es

bien sabido, la verificación de que alguno de estos polimorfismos genotipados sea la mutación causal, requiere nuevos experimentos y análisis. Uno de ellos, consistente en realizar estudios de expresión de los diferentes genotipos, se ha llevado a cabo en el presente trabajo y será discutido posteriormente.

Los resultados obtenidos en el análisis de asociación del índice C20:4/C18:2 indican un efecto significativo del SNP 845 sobre el mismo. Este SNP está localizado en el gen *arachidonate 12-lipoxygenase (ALOX12)* que codifica para una proteína que participa en la ruta de oxidación del ácido araquidónico a ácido linoleico mediante la oxidación del carbono 12 del araquidonato en ácido 12-hidroxicosaenoico. El índice C20:4/C18:2 expresa los coeficientes de actividad de las enzimas de las rutas biosintéticas de conversión de ácido araquidónico a ácido linoleico o viceversa (Pamplona *et al.*, 1998). Muchos mamíferos son incapaces de sintetizar estos ácidos grasos esenciales y deben tomarlos de la dieta. La participación de la proteína codificada por el gen *ALOX12*, relacionado con el índice sobre el que se detectó la asociación, conviértelo en un sólido gen candidato para explicar los efectos sobre este carácter. Al igual que en el caso del *PCTP*, no se puede descartar ni aceptar que este SNP sea la mutación causal responsable del efecto, por lo tanto es recomendable realizar experimentos adicionales para comprobar el efecto en otras poblaciones diferentes, análisis de expresión así como análisis a nivel proteico.

Cabe destacar que en el análisis de detección de QTLs no se detectó ningún efecto sobre el índice C20:4/C18:2. La hipótesis de QTL implica la presencia de alelos alternativos fijados en las líneas parentales, lo que probablemente no se corresponde con la segregación de la mutación causal en el material experimental, que presenta un mayor desequilibrio de ligamiento con los SNP detectados como significativos.

Un último aspecto a tener en cuenta es que, tal como ya se ha mencionado, el chip *SNP60* fue desarrollado a partir de la información de 19 librerías de ADN de individuos Landrace, Large-White, Duroc y Pietrain, con lo cual no se tuvo en cuenta la segregación de estos SNPs en otras razas o líneas genéticas diferentes como es el caso de la raza Ibérica (Ramos *et al.*, 2009), por lo tanto puede suceder que los SNPs utilizados en el análisis de asociación sean menos informativos de lo requerido para este tipo de análisis.

8.5. ANÁLISIS DE EXPRESIÓN

Los datos procedentes de *microarrays* de expresión utilizados en este estudio fueron obtenidos en animales RC2, dentro del plan de trabajo del subproyecto AGL2008-04818-C03-02 del Plan Nacional de Investigación (2008-2011), a cargo del Área de Producción Animal del Centro UdL-IRTA (Lleida, España). El objetivo de este subproyecto es caracterizar a nivel genómico patrones de expresión génica diferencial y de redes génicas implicadas en el crecimiento y la calidad de la carne porcina, así como rutas metabólicas para mejorar estos caracteres modificando uno o varios enlaces de la red. Una parte de nuestros resultados justifican un estudio a nivel de expresión génica en músculo *longissimus*, por lo que la cesión por parte del Área de Producción Animal del Centro UdL-IRTA de los datos de escaneado de los mismos posibilitó llevarlos a cabo.

En este estudio se realizaron cinco análisis de expresión con objetivos diferentes. El primero de ellos consistía en buscar genes diferencialmente expresados entre individuos con distinto genotipo probable del QTL. Como los individuos de los que se dispone de *microarrays* pertenecen a un retrocruce, sólo se pueden detectar individuo *Qq* y *qq*, lo que conlleva una disminución en la potencia del análisis. Al realizar el análisis de expresión las diferencias de expresión de los genes, pueden atribuirse a los efectos del QTL. El propósito de este análisis consistía, principalmente, en detectar genes candidatos que puedan explicar los efectos del QTL. Si los genes diferencialmente expresados están localizados dentro del intervalo de confianza del QTL o próximos al mismo, las diferencias de expresión se pueden atribuir a mutaciones localizadas en las regiones reguladoras de estos mismos genes. Si por el contrario los genes diferencialmente expresados están fuera del intervalo de confianza, las diferencias de expresión se explicarían por cambios conformacionales de una proteína que regule su expresión. Por lo tanto, habría que verificar que la proteína está codificada por un gen localizado dentro del intervalo de confianza del QTL y realizar una búsqueda de polimorfismos en la secuencia codificante de la misma.

Como resultado de este primer análisis, se detectaron tres sondas localizadas en el SSC12 de un total de cuatro con valores de $p < 0.001$ y con un $FDR < 0.10$. La cuarta de las sondas está localizada en el SSC14, y ninguna de estas sondas se corresponde con transcritos de genes que codifiquen proteínas implicadas en el metabolismo de ácidos grasos. De las 18 sondas restantes con un valor de $p < 0.001$ y $0.10 < FDR < 0.15$, dos están localizadas dentro del intervalo de confianza del QTL. Estas sondas corresponden a los genes que codifican las proteínas PCTP y TMEM97. Debido, a su localización en el

intervalo de confianza del QTL, al hecho de que estos genes codifiquen para una proteína con función relacionada con el metabolismo de ácidos grasos y a las considerables diferencias de expresión detectadas (el ratio *Qq/qq* correspondiente a la sonda del gen *PCTP* toma un valor de 0.77 y el de la sonda del gen *TMEM97*, un valor de 1.26) se consideraron estas sondas significativas, a pesar de que el valor calculado para su FDR excede ligeramente el umbral estadístico establecido de 0.10.

Para los restantes genes diferencialmente expresados, incluidos en la tabla 7.1, no se ha encontrado ninguna relación con el metabolismo lipídico en los localizados en el SSC12, por lo que no podemos considerarlos candidatos funcionales. En cuanto a los localizados en otros cromosomas, tampoco se consideran buenos genes candidatos ya que no hay descrita ninguna interacción entre ellos y genes localizados en el SSC12.

En los otros cuatro análisis, se estudiaron los efectos sobre la expresión génica de los SNPs de los genes *ACACA* y *PCTP*, significativamente asociados al contenido de C16:0 en grasa intramuscular. Se observó la expresión diferencial de 20 sondas como efecto del número de copias del alelo A del *SNP636* anotado en el gen *ACACA* (*ASGA0066302* G>A). Las funciones de los genes correspondientes a estas sondas no tienen relación aparente con el metabolismo de ácidos grasos, ni tampoco interacciones conocidas con genes relacionados con el mismo (Tabla 7.2). En el mismo análisis, ninguna de las cinco sondas del gen *ACACA* incluidas en el *array* de Affymetrix presentó diferencias de expresión significativas. Estos resultados, condicionados al SNP analizado, no aportan evidencias favorables a la posible causalidad del gen *ACACA* de los efectos génicos observados sobre el contenido de ácido palmítico en grasa intramuscular.

En los tres análisis restantes se estudió el efecto sobre la expresión de los genotipos de los *SNP542* (*ASGA0054039T>C*) y *SNP544* (*ASGA0054041A>G*) del gen *PCTP*. En dos de ellos se estimó el efecto del número de copias de uno de los alelos de cada polimorfismo (Tablas 7.3 y 7.4) y en el tercero se compararon los tres genotipos del *SNP544* (Tabla 7.5). En todos ellos la sonda con las diferencias de expresión más significativas es la única del *array* de expresión correspondiente al propio gen *PCTP*. Los efectos son particularmente significativos para el *SNP544* (*ASGA0054041A>G*) que presenta en los animales RC2 frecuencias intermedias, más informativas que las del otro polimorfismo.

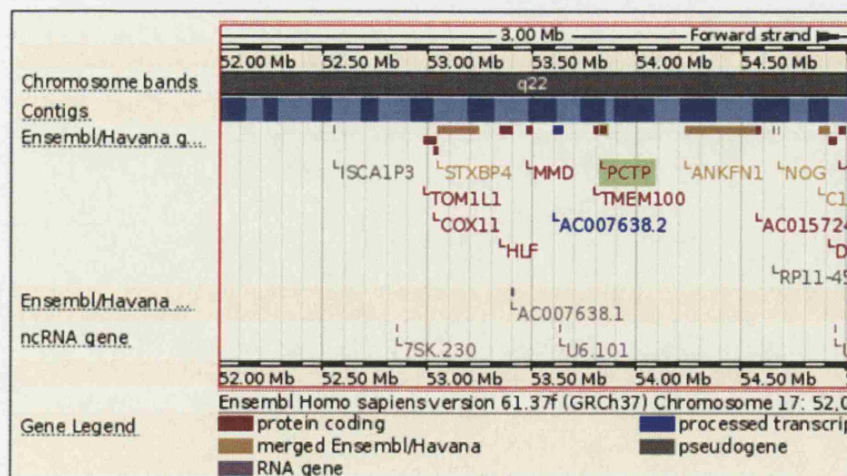
Estos resultados junto el efecto de ambos SNPs sobre C16:0 y además la localización del gen *PCTP* en la posición de 29.65 Mb, próximo al intervalo de confianza del QTL (30.96-35.02 Mb), avalarían la hipótesis de que uno de los dos fuera la mutación

DISCUSIÓN

causal o estuviera muy estrechamente ligado a ella. No obstante, falta por establecer una relación entre el nivel de expresión y contenido de ácido palmítico que requiere disponer de medidas precisas de la primera, obtenidas por PCR cuantitativa. La validación de la hipótesis requiere además realizar otros ensayos. En primer lugar, los SNPs del gen *PCTP* deben ser testados en diferentes poblaciones para comprobar que se mantienen sus efectos sobre el ácido palmítico (C16:0). Por otra parte, sería interesante caracterizar el ADNc de este gen así como identificar SNPs exónicos que pudieran estar relacionados con una alteración de la proteína. Otra tarea prevista será la obtención de registros de composición de ácidos grasos en la grasa subcutánea de estos animales (F3+RC1+RC2), así como de los genotipos de estos polimorfismos en individuos F2, con lo que se podrá comprobar sus posibles efectos sobre la composición lipídica de grasa subcutánea.

Asimismo, en los tres análisis realizados a partir de los genotipos de los SNPs del gen *PCTP* aparece diferencialmente expresado el gen *Cytochrome c oxidase assembly homolog* (*COX11*), representado por dos sondas diferentes (Tablas 7.3, 7.4 y 7.5). Este gen está localizado en la posición de 28.93 Mb del SSC12, adyacente al gen *PCTP*. Esta expresión diferencial podría ser atribuida a que tienen una regulación de la expresión común, y por lo tanto la mutación que está regulando la expresión de *PCTP*, también regula su expresión.

Figura 8.4. Posición de los genes *PCTP*, *COX11* y *HLF* en humano.



Ambos genes *PCTP* y *COX11* ocupan localizaciones cercanas en el genoma humano (cromosoma 17), como se puede observar en la figura 8.4. Entre ellos está situado el gen *Hepatic leukemia factor* (*HLF*) que codifica para el factor de transcripción HLF. Flowers *et al.* (2008) compararon la expresión diferencial de ratones control con ratones deficientes en la

DISCUSIÓN

enzima SCD. Observaron que los segundos tenían reprimida la expresión de genes implicados en detoxificación, biosíntesis, elongación, desaturación, oxidación, transporte y cetogénesis de ácidos grasos. Uno de los genes reprimidos era el gen *HLF*, la coexpresión del mismo junto con genes implicados en el metabolismo de ácidos grasos podría indicar que participa en la regulación transcripcional de alguno de ellos, por lo tanto, este factor de transcripción también podría estar regulando a *PCTP* y ser el portador de la mutación causal. Este y otros estudios futuros quedan sin embargo fuera del horizonte de esta tesis doctoral.

CONCLUSIONES

CONCLUSIONES

1. Se ha construido un mapa de ligamiento del SSC12 a partir de la información de 978 SNPs del *PorcineSNP60 Genotyping Bead Chips* anotados en este cromosoma. Se han detectado valores elevados de las frecuencias de recombinación de las sondas *ASGA0104321* y *MARC0110971* que indican altamente improbable su localización en este cromosoma y alertan para la existencia de errores en el mapa físico porcino.
2. A partir de la información proporcionada por el *PorcineSNP60 Genotyping Bead Chip* y de los mapas de ligamiento desarrollados en este estudio se realizaron dos estimas de la relación entre las distancias genéticas y el número de Megabases, que exceden la convencionalmente aceptada (1Mb = 1 cM), tanto en el genoma humano como en los de los animales domésticos.
3. En el presente trabajo se caracterizó el ADNc de tres genes candidatos (*ACOX*, *ACLY* y *SREBF1*) localizados en el SSC12 en animales pertenecientes a las razas Ibérico y Landrace. El análisis de asociación de los polimorfismos identificados en estos genes con el contenido de ácidos grasos de la grasa dorsal no ha generado resultados significativos.
4. Se ha confirmado la existencia de un QTL, en el SSC12, con efecto sobre los porcentajes de los ácidos palmitoleico [C16:1(n-9)], esteárico (C18:0) y vaccénico [C18:1(n-7)] sobre el total de ácidos grasos presentes en la grasa dorsal. Asimismo se han confirmado los resultados significativos de la asociación entre el SNP *ACACA*: c.5634T>C y los porcentajes de estos ácidos grasos. En una región intrónica del gen *ACACA* con homología al Promotor I (PI) bovino, se ha identificado un polimorfismo (*ACACA*: g.48001A>G) con efecto sobre el porcentaje de ácido esteárico (C18:0) de la grasa dorsal.
5. Se ha detectado un QTL, en el SSC12, con efecto sobre el porcentaje de ácido palmítico (C16:0) en la grasa intramuscular. Este resultado se ha obtenido consistentemente en análisis realizados con un número creciente de marcadores (hasta 489 SNPs). Los efectos detectados sobre el mismo son de una magnitud del 0.6 % por copia del alelo *Q* Ibérico, equivalentes a un 2.6% de la media del carácter.
6. En el análisis de asociación llevado a cabo con los datos de genotipado del *PorcineSNP60 Genotyping Bead Chip* se han detectado efectos significativos de sondas localizadas en los genes *ACACA* y *PCTP*, con función relacionada con el metabolismo de ácidos grasos, sobre el contenido en ácido palmítico (C16:0) de la grasa intramuscular. Estos resultados sugieren el gen *PCTP* como buen candidato funcional y posicional.

CONCLUSIONES

7. Los resultados del análisis de asociación indican asimismo un efecto significativo sobre el cociente entre los ácidos araquidónico y linoleico (C20:4/C18:2) de un SNP en el gen *ALOX12* y proporcionan evidencia en favor de su estudio como gen candidato.
8. Se ha estudiado en animales RC2, el efecto del genotipo del QTL para ácido palmítico (C16:0) sobre la expresión génica en músculo *longissimus*, utilizando el *array* de expresión de Affymetrix. Los resultados evidencian diferencias de expresión en dos genes, *PCTP* y *TMEM97*, con cocientes de expresión entre los genotipos *Qq/qq* de 0.77 y 1.26, respectivamente. Ambos genes están localizados en el intervalo de confianza del QTL e implicados en el metabolismo lipídico.
9. Se observaron efectos altamente significativos de los genotipos de los SNPs *ASGA0054039T>C* y *ASGA0054041A>G* del gen *PCTP* sobre la expresión del propio gen, por lo que se le considera como el principal candidato posicional y funcional subyacente al QTL para contenido en ácido palmítico de la grasa intramuscular.

CONCLUSIONS

CONCLUSIONS

1. A new linkage map for SSC12 has been built using 978 SNPs from the *PorcineSNP60 Genotyping Bead Chips* annotated in this chromosome. High recombination frequencies have been detected for the probes *ASGA0104321* and *MARC0110971*, which probably indicates that there are not located in this chromosome and point out the existence of mistakes in the porcine physical map.
2. Two estimations of the equivalence between genetic distances and the Megabase number using the information from the *PorcineSNP60 Genotyping Bead Chips* and the linkage maps were conducted, both estimates were superior to the conventionally assumed (1 Mb = 1 cM) for the human genome as well as in domestic animals.
3. The cDNA of three candidate genes (*ACOX*, *ACLY* and *SREBF1*) mapped on SSC12 has been characterized in purebred Iberian and Landrace animals. The association analyses conducted for the identified polymorphisms in these genes and the backfat fatty acid composition did not revealed significant results.
4. A QTL in SSC12 has been confirmed affecting the percentages of palmitoleic [C16:1(n-9)], stearic (C18:0) and vaccenic [C18:1(n-7)] fatty acids measured on the total of backfat fatty acid. In addition, the significant associations of the SNP *ACACA*: c.5634T>C with fatty acid percentages have been also confirmed. A new polymorphism (*ACACA*: g.48001A>G) affecting the backfat stearic fatty acid (C18:0) percentage has been identified in an intronic region of *ACACA*, highly similar to the bovine promoter I (PI).
5. A QTL in SSC12 affecting the palmitic fatty acid (C16:0) percentage in intramuscular fat has been detected. This result has been consistently obtained in the analyses conducted while increasing the number of markers (up to 489 SNPs). The magnitude of the effect detected on the trait is 0.6% for the *Q* Iberian allele, equivalent to 2.6% of the mean of the trait.
6. In the association analysis performed with the genotyping data from the *PorcineSNP60 Genotyping Bead Chip*, significant effects on the palmitic fatty acid (C16:0) percentage in intramuscular fat were detected in the *ACACA* and *PCTP* genes probes, both with functions related with the fatty acid metabolism,. These results suggest that *PCTP* is a good functional and positional candidate gene.

CONCLUSIONS

7. In addition, the association analysis results show a significant effect on the ratio between araquidonic and linoleic fatty acids (C20:4/C18:2) of a SNP in the *ALOX12* gene, providing evidences of its relevance as candidate gene.
8. The effect on the *longissimus* muscle gene expression of the QTL genotype for palmitic fatty acid has been investigated in RC2 animals using the Affymetrix gene expression microarrays. The results show expression differences in two genes, *PCTP* and *TMEM97* and expression ratios between the Qq/qq genotypes were 0.77 and 1.26, respectively. Both genes are involved in the lipid metabolism and mapped within the confidence interval of the QTL.
9. The SNPs *ASGA0054039T>C* and *ASGA0054041A>G* of the *PCTP* gene showed highly significant effects on its own gene expression, therefore *PTCP* is considered the main positional and functional candidate underlying the intramuscular fat QTL for palmitic fatty acid.

BIBLIOGRAFÍA

- Andersson, L. (2001). *Genetic dissection of phenotypic diversity in farm animals*. Nat Rev Genet **2**(2): 130-138.
- Aulchenko, Y. S., D. J. de Koning & C. Haley (2007). *Genomewide rapid association using mixed model and regression: A fast and simple method for genomewide pedigree-based quantitative trait loci association analysis*. Genetics **177**: 577-585.
- Barton, L., T. Kott, D. Bures, D. Rehak, R. Zahradkova & B. Kottova (2010). *The polymorphisms of stearoyl-CoA desaturase (SCD1) and sterol regulatory element binding protein-1 (SREBP-1) genes and their association with the fatty acid profile of muscle and subcutaneous fat in Fleckvieh bulls*. Meat Science **85**(1): 15-20.
- Benjamini, Y. & Y. Hochberg (1995). *Controlling the false discovery rate: a practical and powerful approach to multiple testing*. Journal of the Royal Statistical Society, Journal of the Royal Statistical Society, Series B (Methodological) **57**(1): 289-300.
- Benjamini, Y. & D. Yekutieli (2005). *Quantitative trait Loci analysis using the false discovery rate*. Genetics **171**(2): 783-790.
- Bhuiyan, M. S. A., S. L. Yu, J. T. Jeon, D. Yoon, Y. M. Cho, E. W. Park, N. K. Kim, K. S. Kim & J. H. Lee (2009). *DNA Polymorphisms in SREBF1 and FASN Genes Affect Fatty Acid Composition in Korean Cattle (Hanwoo)*. Asian-Australasian Journal of Animal Sciences **22**(6): 765-773.
- Braglia, S., O. Ramirez, J. L. Noguera, A. Tomas, C. Ovilo & L. Varona (2006). *Comparison of genetic models for analysing the effects of a PvuII polymorphism in the oestrogen receptor 1 (ESR1) gene on prolificacy in an Iberian x Meishan pig population*. Anim Genet **37**(5): 454-458.
- Cartharius, K., K. Frech, K. Grote, B. Klocke, M. Haltmeier, A. Klingenhoff, M. Frisch, M. Bayerlein & T. Werner (2005). *MatInspector and beyond: promoter analysis based on transcription factor binding sites*. Bioinformatics **21**(13): 2933-2942.
- Casellas, J., N. Ibanez-Escriche, M. Martinez-Giner & L. Varona (2008). *GEAMM v.1.4: a versatile program for mixed model analysis of gene expression data*. Anim Genet **39**(1): 89-90.
- Cava, R., J. Ruiz, J. Ventanas & T. Antequera (1999). *Effect of alpha-tocopheryl acetate supplementation and the extensive feeding of pigs on the volatile aldehydes during the maturation of Iberian ham*. Food Science and Technology International **5**(3): 235-241.

BIBLIOGRAFÍA

- Chen, J., X. J. Yang, D. Xia, J. Wegner, Z. Jiang & R. Q. Zhao (2008). *Sterol regulatory element binding transcription factor 1 expression and genetic polymorphism significantly affect intramuscular fat deposition in the longissimus muscle of Erhualian and Suta pigs*. J Anim Sci **86**(1): 57-63.
- Clop, A., C. Ovilo, M. Perez-Enciso, A. Cercos, A. Tomas, A. Fernandez, A. Coll, J. M. Folch, C. Barragan, et al. (2003). *Detection of QTL affecting fatty acid composition in the pig*. Mamm Genome **14**(9): 650-656.
- Daval, S., S. Lagarrigue & M. Douaire (2000). *Messenger RNA levels and transcription rates of hepatic lipogenesis genes in genetically lean and fat chickens*. Genet Sel Evol **32**(5): 521-531.
- Dmochowska, A., D. Dignard, R. Maleszka & D. Y. Thomas (1990). *Structure and transcriptional control of the Saccharomyces cerevisiae POX1 gene encoding acyl-coenzyme A oxidase*. Gene **88**(2): 247-252.
- Duijvesteijn, N., E. F. Knol, J. W. Merks, R. P. Crooijmans, M. A. Groenen, H. Bovenhuis & B. Harlizius (2010). *A genome-wide association study on androstene levels in pigs reveals a cluster of candidate genes on chromosome 6*. BMC Genet **11**: 42.
- Elshourbagy, N. A., J. C. Near, P. J. Kmetz, T. N. Wells, P. H. Groot, B. A. Saxty, S. A. Hughes, M. Franklin & I. S. Gloger (1992). *Cloning and expression of a human ATP-citrate lyase cDNA*. Eur J Biochem **204**(2): 491-499.
- Erbe, M., F. Ytournal, E. C. G. Pimentel, A. R. Sharifi & H. Simianer (2011). *Power and robustness of the whole genome association mapping approaches in selected populations*. Journ Anim Breed Genet **128**: 3-14.
- Fan, B., Z. Q. Du, D. M. Gorbach & R. M. F. (2010). *Development and Application of High-density SNP Arrays in Genomic Studies of Domestic Animals*. Asian-Aust. J. Anim. Sci. **23**(7): 833-847.
- Fan, C. Y., J. Pan, R. Chu, D. Lee, K. D. Kluckman, N. Usuda, I. Singh, A. V. Yeldandi, M. S. Rao, et al. (1996). *Hepatocellular and hepatic peroxisomal alterations in mice with a disrupted peroxisomal fatty acyl-coenzyme A oxidase gene*. J Biol Chem **271**(40): 24698-24710.
- Fernández, A., E. de Pedro, N. Nunez, L. Silio, J. Garcia-Casco & C. Rodriguez (2003). *Genetic parameters for meat and fat quality and carcass composition traits in Iberian pigs*. Meat Science **64**(4): 405-410.

BIBLIOGRAFÍA

- Fernández, A. I., E. Alves, Y. Ramayo-Caldas, A. Mercadé, J. L. Noguera, L. Silió, J. M. Folch & C. Rodríguez (2011). *Detección clásica de QTL utilizando nuevas herramientas genómicas XIV Jornadas sobre producción animal*. A. I. p. e. D. Agrario. Zaragoza, Sin publicar.
- Flowers, M. T., M. P. Keller, Y. Choi, H. Lan, C. Kendziorski, J. M. Ntambi & A. D. Attie (2008). *Liver gene expression analysis reveals endoplasmic reticulum stress and metabolic dysfunction in SCD1-deficient mice fed a very low-fat diet*. *Physiol Genomics* **33**(3): 361-372.
- Foretz, M., C. Pacot, I. Dugail, P. Lemarchand, C. Guichard, X. Le Liepvre, C. Berthelie-Lubrano, B. Spiegelman, J. B. Kim, et al. (1999). *ADD1/SREBP-1c is required in the activation of hepatic lipogenic gene expression by glucose*. *Mol Cell Biol* **19**(5): 3760-3768.
- Fournier, B., J. M. Saudubray, B. Benichou, S. Lyonnet, A. Munnich, H. Clevers & B. T. Poll-The (1994). *Large deletion of the peroxisomal acyl-CoA oxidase gene in pseudoneonatal adrenoleukodystrophy*. *J Clin Invest* **94**(2): 526-531.
- Friedman, R. C., K. K. Farh, C. B. Burge & D. P. Bartel (2009). *Most mammalian mRNAs are conserved targets of microRNAs*. *Genome Res* **19**(1): 92-105.
- Gallardo, D., R. Quintanilla, L. Varona, I. Diaz, O. Ramirez, R. N. Pena & M. Amills (2009). *Polymorphism of the pig acetyl-coenzyme A carboxylase alpha gene is associated with fatty acid composition in a Duroc commercial line*. *Anim Genet* **40**(4): 410-417.
- Gao, X., L. C. Becker, D. M. Becker, J. D. Starmer & M. A. Province (2010). *Avoiding the high Bonferroni penalty in genome-wide association studies*. *Genet Epidemiol* **34**(1): 100-105.
- Gibson, D. M., R. T. Lyons, D. F. Scott & Y. Muto (1972). *Synthesis and degradation of the lipogenic enzymes of rat liver*. *Adv Enzyme Regul* **10**: 187-204.
- Goddard, M. E. & B. J. Hayes (2009). *Mapping genes for complex traits in domestic animals and their use in breeding programmes*. *Nat Rev Genet* **10**(6): 381-391.
- Gorbach, D., B. Mote, L. Totir, R. Fernando & M. Rothschild (2010). *Polydactyl inheritance in the pig*. *J Hered* **101**(4): 469-475.
- Green, P., K. Falls & S. Crooks (1990). *Documentation for Crimap*. Unpublished mimeo. Available: <http://linkage.Rockefeller.Edu/soft/crimap>.

- Hayes, B. J., P. M. Visscher, H. C. McPartlan & M. E. Goddard (2003). *Novel multilocus measure of linkage disequilibrium to estimate past effective population size*. *Genome Res* **13**(4): 635-643.
- Heinemeyer, T., E. Wingender, I. Reuter, H. Hermjakob, A. E. Kel, O. V. Kel, E. V. Ignatieva, E. A. Ananko, O. A. Podkolodnaya, et al. (1998). *Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL*. *Nucleic Acids Res* **26**(1): 362-367.
- Hoashi, S., N. Ashida, H. Ohsaki, T. Utsugi, S. Sasazaki, M. Taniguchi, K. Oyama, F. Mukai & H. Mannen (2007). *Genotype of bovine sterol regulatory element binding protein-1 (SREBP-1) is associated with fatty acid composition in Japanese Black cattle*. *Mammalian Genome* **18**(12): 880-886.
- Huang da, W., B. T. Sherman & R. A. Lempicki (2009). *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources*. *Nat Protoc* **4**(1): 44-57.
- Jimenez-Colmenero, F., J. Ventanas & F. Toldra (2010). *Nutritional composition of dry-cured ham and its role in a healthy diet*. *Meat Science* **84**(4): 585-593.
- Kang, H. W., K. Kanno, E. F. Scapa & D. E. Cohen (2010a). *Regulatory role for phosphatidylcholine transfer protein/StarD2 in the metabolic response to peroxisome proliferator activated receptor alpha (PPARalpha)*. *Biochim Biophys Acta* **1801**(4): 496-502.
- Kang, H. W., J. Wei & D. E. Cohen (2010b). *PC-TP/StarD2: Of membranes and metabolism*. *Trends Endocrinol Metab* **21**(7): 449-456.
- Kanno, K., M. K. Wu, E. F. Scapa, S. L. Roderick & D. E. Cohen (2007). *Structure and function of phosphatidylcholine transfer protein (PC-TP)/StarD2*. *Biochim Biophys Acta* **1771**(6): 654-662.
- Katsurada, A., H. Fukuda & N. Iritani (1986). *Effects of dietary nutrients on substrate and effector levels of lipogenic enzymes, and lipogenesis from tritiated water in rat liver*. *Biochim Biophys Acta* **878**(2): 200-208.
- Kim, Y., M. Kong, Y. J. Nam & C. Lee (2006). *A quantitative trait locus for oleic fatty acid content on Sus scrofa chromosome 7*. *Journal of Heredity* **97**(5): 535-537.

BIBLIOGRAFÍA

- Knowles, S. E., I. G. Jarrett, O. H. Filsell & F. J. Ballard (1974). *Production and utilization of acetate in mammals*. Biochem J **142**(2): 401-411.
- Krauss, R. M., R. H. Eckel, B. Howard, L. J. Appel, S. R. Daniels, R. J. Deckelbaum, J. W. Erdman, Jr., P. Kris-Etherton, I. J. Goldberg, et al. (2000). *AHA Dietary Guidelines: revision 2000: A statement for healthcare professionals from the Nutrition Committee of the American Heart Association*. Circulation **102**(18): 2284-2299.
- Krek, A., D. Grun, M. N. Poy, R. Wolf, L. Rosenberg, E. J. Epstein, P. MacMenamin, I. da Piedade, K. C. Gunsalus, et al. (2005). *Combinatorial microRNA target predictions*. Nat Genet **37**(5): 495-500.
- Kunau, W. H., V. Dommès & H. Schulz (1995). *beta-oxidation of fatty acids in mitochondria, peroxisomes, and bacteria: a century of continued progress*. Prog Lipid Res **34**(4): 267-342.
- Lander, E. & L. Kruglyak (1995). *Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results*. Nat Genet **11**(3): 241-247.
- Lantieri, F., M. A. Jhun, J. Park, T. Park & M. Devoto (2009). *Comparative analysis of different approaches for dealing with candidate regions in the context of a genome-wide association study*. BMC Proc **3 Suppl 7**: S93.
- Lawrence, T. J. L. & V. R. Fowler (1997). *Growth of animals*. New York, NY, CAB International.
- Ledur, M. C., N. Navarro & M. Perez-Enciso (2010). *Large-scale SNP genotyping in crosses between outbred lines: how useful is it?* Heredity **105**(2): 173-182.
- Letunic, I., R. R. Copley, B. Pils, S. Pinkert, J. Schultz & P. Bork (2006). *SMART 5: domains in the context of genomes and networks*. Nucleic Acids Res **34**(Database issue): D257-260.
- Li, H. D., M. S. Lund, O. F. Christensen, V. R. Gregersen, P. Henckel & C. Bendixen (2010). *Quantitative trait loci analysis of swine meat quality traits*. J Anim Sci **88**(9): 2904-2912.
- Lichtenstein, A. H. (2003). *Dietary fat and cardiovascular disease risk: quantity or quality?* J Womens Health (Larchmt) **12**(2): 109-114.
- Lin, J., S. M. Zhang, N. R. Cook, I. M. Lee & J. E. Buring (2004). *Dietary fat and fatty acids and risk of colorectal cancer in women*. Am J Epidemiol **160**(10): 1011-1022.

BIBLIOGRAFÍA

- Livak, K. J. & T. D. Schmittgen (2001). *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method*. Methods **25**(4): 402-408.
- lo Fiego, D. P., P. Santoro, P. Macchioni & E. de Leonibus (2005). *Influence of genetic type, live weight at slaughter and carcass fatness on fatty acid composition of subcutaneous adipose tissue of raw ham in the heavy pig*. Meat Science **69**(1): 69 (61) 107-114.
- Lonergan, S. M., E. Huff-Lonergan, L. J. Rowe, D. L. Kuhlers & S. B. Jungst (2001). *Selection for lean growth efficiency in Duroc pigs influences pork quality*. J Anim Sci **79**(8): 2075-2085.
- Lopez-Bote, C. J. (1998). *Sustained utilization of the Iberian pig breed*. Meat Science **49**(Suppl. 1): 49 (Suppl. 41) S17-S27.
- Lopez, M. O., L. Delahoz, M. I. Cambero, E. Gallardo, G. Reglero & J. A. Ordóñez (1992). *VOLATILE COMPOUNDS OF DRY HAM FROM IBERIAN PIGS*. Meat Science **31**(3): 267-277.
- Mangin, B., B. Goffinet & A. Rebai (1994). *Constructing confidence intervals for QTL location*. Genetics **138**(4): 1301-1308.
- Mao, J., S. S. Chirala & S. J. Wakil (2003). *Human acetyl-CoA carboxylase 1 gene: presence of three promoters and heterogeneity at the 5'-untranslated mRNA region*. Proc Natl Acad Sci U S A **100**(13): 7515-7520.
- Mao, J., S. Marcos, S. K. Davis, J. Burzlaff & H. M. Seyfert (2001). *Genomic distribution of three promoters of the bovine gene encoding acetyl-CoA carboxylase alpha and evidence that the nutritionally regulated promoter I contains a repressive element different from that in rat*. Biochem J **358**(Pt 1): 127-135.
- Mente, A., L. de Koning, H. S. Shannon & S. S. Anand (2009). *A Systematic Review of the Evidence Supporting a Causal Link Between Dietary Factors and Coronary Heart Disease*. Archives of Internal Medicine **169**(7): 659-669.
- Metzker, M. L. (2010). *Sequencing technologies - the next generation*. Nat Rev Genet **11**(1): 31-46.
- Meuwissen, T. H., B. J. Hayes & M. E. Goddard (2001). *Prediction of total genetic value using genome-wide dense marker maps*. Genetics **157**(4): 1819-1829.

- Miyazawa, S., H. Hayashi, M. Hijikata, N. Ishii, S. Furuta, H. Kagamiyama, T. Osumi & T. Hashimoto (1987). *Complete nucleotide sequence of cDNA and predicted amino acid sequence of rat acyl-CoA oxidase*. J Biol Chem **262**(17): 8131-8137.
- Moon, Y. A., K. S. Kim, S. W. Park & Y. S. Kim (1996). *Cloning and identification of exon-intron organization of the rat ATP-citrate lyase gene*. Biochimica Et Biophysica Acta-Gene Structure and Expression **1307**(3): 280-284.
- Morais, S., A. Knoll-Gellida, M. Andre, C. Barthe & P. J. Babin (2007). *Conserved expression of alternative splicing variants of peroxisomal acyl-CoA oxidase 1 in vertebrates and developmental and nutritional regulation in fish*. Physiol Genomics **28**(3): 239-252.
- Muñoz, G., E. Alcazar, A. Fernández, C. Barragán, A. Carrasco, E. de Pedro, L. Silió, J. L. Sánchez & M. C. Rodríguez (2010). *Effects of porcine MC4R and LEPR polymorphisms, gender and Duroc sire line on economic traits in Duroc × Iberian crossbred pigs*. Meat Sci doi **10.1016/j.meatsci.2010.12.018**.
- Muñoz, G., E. Alves, A. Fernandez, C. Ovilo, C. Barragan, J. Estelle, R. Quintanilla, J. M. Folch, L. Silió, et al. (2007). *QTL detection on porcine chromosome 12 for fatty-acid composition and association analyses of the fatty acid synthase, gastric inhibitory polypeptide and acetyl-coenzyme A carboxylase alpha genes*. Anim Genet **38**(6): 639-646.
- Nezer, C., L. Moreau, D. Wagenaar & M. Georges (2002). *Results of a whole genome scan targeting QTL for growth and carcass traits in a Pietrain × Large White intercross*. Genet Sel Evol **34**(3): 371-387.
- Ng, P. C. & S. Henikoff (2006). *Predicting the effects of amino acid substitutions on protein function*. Annu Rev Genomics Hum Genet **7**: 61-80.
- Ngo, S. N., R. A. McKinnon & I. Stupans (2003). *Identification and cloning of two forms of liver peroxisomal fatty Acyl CoA Oxidase from the koala (Phascolarctos cinereus)*. Gene **309**(2): 91-99.
- Nii, M., T. Hayashi, F. Tani, A. Niki, N. Mori, N. Fujishima-Kanaya, M. Komatsu, K. Aikawa, T. Awata, et al. (2006). *Quantitative trait loci mapping for fatty acid composition traits in perirenal and back fat using a Japanese wild boar × Large White intercross*. Animal Genetics **37**(4): 342-347.

BIBLIOGRAFÍA

- Nohammer, C., Y. El-Shabrawi, S. Schauer, M. Hiden, J. Berger, S. Forss-Petter, E. Winter, R. Eferl, R. Zechner, et al. (2000). *cDNA cloning and analysis of tissue-specific expression of mouse peroxisomal straight-chain acyl-CoA oxidase*. Eur J Biochem **267**(4): 1254-1260.
- Ohsaki, H., A. Tanaka, S. Hoashi, S. Sasazaki, K. Oyama, M. Taniguchi, F. Mukai & H. Mannen (2009). *Effect of SCD and SREBP genotypes on fatty acid composition in adipose tissue of Japanese Black cattle herds*. Anim Sci J **80**(3): 225-232.
- Osumi, T., N. Ishii, S. Miyazawa & T. Hashimoto (1987). *Isolation and structural characterization of the rat acyl-CoA oxidase gene*. J Biol Chem **262**(17): 8138-8143.
- Óvilo, C., A. Fernández, A. I. Fernández, P. M. Palomino, J. Rodrigañez, C. Rodríguez, L. Silio & C. Lopez-Bote (2010). *Effect of dietary oleic acid content: Different genetic regulation of fatty acid metabolism on muscle and fat of Iberian pigs*. VII Congreso Internacional de Cerdo Mediterráneo. Córdoba.
- Pamplona, R., M. Portero-Otin, D. Riba, C. Ruiz, J. Prat, M. J. Bellmunt & G. Barja (1998). *Mitochondrial membrane peroxidizability index is inversely related to maximum life span in mammals*. J Lipid Res **39**(10): 1989-1994.
- Pan, H. J., D. S. Agate, B. L. King, M. K. Wu, S. L. Roderick, E. H. Leiter & D. E. Cohen (2006). *A polymorphism in New Zealand inbred mouse strains that inactivates phosphatidylcholine transfer protein*. FEBS Lett **580**(25): 5953-5958.
- Perez-Enciso, M., A. Clop, J. L. Noguera, C. Ovilo, A. Coll, J. M. Folch, D. Babot, J. Estany, M. A. Oliver, et al. (2000). *A QTL on pig chromosome 4 affects fatty acid metabolism: evidence from an Iberian by Landrace intercross*. J Anim Sci **78**(10): 2525-2531.
- Perez-Enciso, M. & I. Misztal (2004). *Qxpak: a versatile mixed model application for genetical genomics and QTL analyses*. Bioinformatics **20**(16): 2792-2798.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. de Bakker, et al. (2007). *PLINK: a tool set for whole-genome association and population-based linkage analyses*. Am J Hum Genet **81**(3): 559-575.
- Ramayo-Caldas, Y., A. Castello, R. N. Pena, E. Alves, A. Mercade, C. A. Souza, A. I. Fernandez, M. Perez-Enciso & J. M. Folch (2010). *Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip*. BMC Genomics **11**: 593.

BIBLIOGRAFÍA

- Ramos, A. M., R. P. Crooijmans, N. A. Affara, A. J. Amaral, A. L. Archibald, J. E. Beever, C. Bendixen, C. Churcher, R. Clark, et al. (2009). *Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology*. PLoS One **4**(8): e6524.
- Reimers, M. & V. J. Carey (2006). *Bioconductor: an open source framework for bioinformatics and computational biology*. Methods Enzymol **411**: 119-134.
- Ren, Z. Q., Y. Wang, Y. J. Xu, L. J. Wang, M. G. Lei, B. Zuo, F. E. Li, D. Q. Xu, R. Zheng, et al. (2008). *Identification of a differentially expressed gene, ACL, between Meishan \times Large White and Large White \times Meishan F1 hybrids and their parents*. Genet Sel Evol **40**(6): 625-637.
- Ronaghi, M. (2003). *Pyrosequencing for SNP genotyping*. Methods Mol Biol **212**: 189-195.
- Ruiz, J., C. Garcia, M. D. Diaz, R. Cava, J. F. Tejeda & J. Ventanas (1999). *Dry cured Iberian ham non-volatile components as affected by the length of the curing process*. Food Research International **32**(9): 643-651.
- Sambrook, J., E. F. Fritsh & T. Maniatis (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, Cold Spring Harbor Laboratory Press.
- Sato, R., A. Okamoto, J. Inoue, W. Miyamoto, Y. Sakai, N. Emoto, H. Shimano & M. Maeda (2000). *Transcriptional regulation of the ATP citrate-lyase gene by sterol regulatory element-binding proteins*. J Biol Chem **275**(17): 12497-12502.
- Sellier, P. & G. Monin (1994). *Genetics of pig meat quality: a review*. Journal of Muscle Foods **5**(2): 5 (2) 187-219.
- Serra, X., F. Gil, M. Pérez-Enciso, M. A. Oliver, J. M. Vázquez, M. Gispert, I. Díaz, F. Moreno, R. Latorre, et al. (1998). *A comparison of carcass, meat quality and histochemical characteristics of Iberian (Guadyerbas line) and Landrace pigs*. Livest Prod Sci **56**: 215-223.
- Simon-Sanchez, J., C. Schulte, J. M. Bras, M. Sharma, J. R. Gibbs, D. Berg, C. Paisan-Ruiz, P. Lichtner, S. W. Scholz, et al. (2009). *Genome-wide association study reveals genetic risk underlying Parkinson's disease*. Nat Genet **41**(12): 1308-1312.
- Stoeckert, C. J., Jr., H. C. Causton & C. A. Ball (2002). *Microarray databases: standards and ontologies*. Nat Genet **32 Suppl**: 469-473.

BIBLIOGRAFÍA

- Storey, J. D. & R. Tibshirani (2003). *Statistical significance for genomewide studies*. Proc Natl Acad Sci U S A **100**(16): 9440-9445.
- Suzuki, M. & H. Okuda (1981). *ATP CITRATE LYASE IN HUMAN ADIPOSE-TISSUE*. Journal of Nutritional Science and Vitaminology **27**(6): 595-598.
- Tenesa, A., P. Navarro, B. J. Hayes, D. L. Duffy, G. M. Clarke, M. E. Goddard & P. M. Visscher (2007). *Recent human effective population size estimated from linkage disequilibrium*. Genome Res **17**(4): 520-526.
- Toro, M., J. Rodríguez, L. Silio & C. Rodríguez (2000). *Genealogical analysis of a closed herd of black hairless Iberian pigs*. Conserv Biol **14**: 1843-1851.
- Torto, B., R. T. Arbogast, D. Van Engelsdorp, S. Willms, D. Purcell, D. Boucias, J. H. Tumlinson & P. E. Teal (2007). *Trapping of Aethina tumida Murray (Coleoptera: Nitidulidae) from Apis mellifera L. (Hymenoptera: Apidae) colonies with an in-hive baited trap*. Environ Entomol **36**(5): 1018-1024.
- Tsai, S., J. P. Cassidy, B. A. Freking, D. J. Nonneman, G. A. Rohrer & J. A. Piedrahita (2006). *Annotation of the Affymetrix porcine genome microarray*. Anim Genet **37**(4): 423-424.
- Uemoto, Y., T. Abe, N. Tameoka, H. Hasebe, K. Inoue, H. Nakajima, N. Shoji, M. Kobayashi & E. Kobayashi (2010). *Whole-genome association study for fatty acid composition of oleic acid in Japanese Black cattle*. Anim Genet.
- Valsta, L. M., H. Tapanainen & S. Mannisto (2005). *Meat fats in nutrition*. Meat Science **70**(3): 525-530.
- Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe & F. Speleman (2002). *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes*. Genome Biol **3**(7): RESEARCH0034.
- Varona, L., L. Gomez-Raya, W. M. Rauw & J. L. Noguera (2005). *A simulation study on the detection of causal mutations from F2 experiments*. J Anim Breed Genet **122**(1): 30-36.
- Ventanas, S., J. Ventanas, J. Tovar, C. Garcia & M. Estevez (2007). *Extensive feeding versus oleic acid and tocopherol enriched mixed diets for the production of Iberian dry-cured hams: Effect on chemical composition, oxidative status and sensory traits*. Meat Science **77**: 246-256.

BIBLIOGRAFÍA

- Wanders, R. J., P. Vreken, S. Ferdinandusse, G. A. Jansen, H. R. Waterham, C. W. van Roermund & E. G. Van Grunsven (2001). *Peroxisomal fatty acid alpha- and beta-oxidation in humans: enzymology, peroxisomal metabolite transporters and peroxisomal diseases*. *Biochem Soc Trans* **29**(Pt 2): 250-267.
- Webb, E. C. & H. A. O'Neill (2008). *The animal fat paradox and meat quality*. *Meat Science* **80**(1): 28-36.
- Wirtz, K. W. (1991). *Phospholipid transfer proteins*. *Annu Rev Biochem* **60**: 73-99.
- Wong, M. L. & J. F. Medrano (2005). *Real-time PCR for mRNA quantitation*. *Biotechniques* **39**(1): 75-85.
- Zhao, H., M. F. Rothschild, R. L. Fernando & J. C. Dekkers (2003). *Tests of candidate genes in breed cross populations for QTL mapping in livestock*. *Mamm Genome* **14**(7): 472-482.
- Zuo, B., H. Yang, M. G. Lei, F. E. Li, C. Y. Deng, C. H. Jiang & Y. Z. Xiong (2007). *Association of the polymorphism in GYS1 and ACOX1 genes with meat quality traits in pigs*. *Animal* **1**(9): 1243-1248.

ANEJO

—

—

ANEJO

Anejo: Sondas utilizadas en la construcción de los dos mapas genéticos del SSC12 utilizados en los diferentes análisis en grasa intramuscular. En la tabla se indica el número de orden atribuido a la sonda, su nombre, la posición física (Mb) y la posición en los mapas de ligamiento 3 (cM¹) y 4 (cM³).

Nº SNP	Sonda	Mb	Posición	
			Mapa 2 (cM)	Mapa 3 (cM)
SNP1	<i>ALGA0124548</i>	0.07	-	0.00
SNP41	<i>ASGA0089708</i>	2.12	-	1.80
SNP51	<i>ASGA0052576</i>	3.56	-	7.40
SNP61	<i>ASGA0052621</i>	4.55	-	8.80
SNP71	<i>H3GA0055229</i>	4.91	-	9.70
SNP91	<i>ASGA0092609</i>	5.51	-	11.20
SNP101	<i>H3GA0033256</i>	5.94	-	12.70
SNP119	<i>MARC0011974</i>	6.7	7.00	-
SNP141	<i>ALGA0064800</i>	7.62	-	16.00
SNP151	<i>DRGA0011583</i>	7.89	-	19.40
SNP161	<i>ASGA0053064</i>	8.2	-	20.40
SNP171	<i>ALGA0106208</i>	8.79	-	21.40
SNP191	<i>MARC0057123</i>	9.76	-	22.10
SNP196	<i>DBWU0000115</i>	9.96	15.60	-
SNP211	<i>SIRI0000304</i>	10.69	-	26.10
SNP231	<i>H3GA0055422</i>	11.55	-	27.80
SNP251	<i>M1GA0016286</i>	12.54	-	28.40

ANEJO

<i>SNP273</i>	<i>ALGA0103808</i>	13.32	21.60	-
<i>SNP301</i>	<i>H3GA0033690</i>	14.54	-	29.80
<i>SNP321</i>	<i>MARC0001291</i>	15.86	-	31.10
<i>SNP350</i>	<i>H3GA0033801</i>	16.99	24.60	-
<i>SNP362</i>	<i>ASGA0082772</i>	17.35	25.70	-
<i>SNP370</i>	<i>H3GA0056346</i>	17.91	27.00	-
<i>SNP371</i>	<i>INRA0038774</i>	18.19	27.10	33.70
<i>SNP377</i>	<i>ALGA0113902</i>	18.92	28.50	-
<i>SNP381</i>	<i>DLAS0001511</i>	19.09	-	35.90
<i>SNP382</i>	<i>SIRI0000605</i>	19.18	32.50	-
<i>SNP391</i>	<i>MARC0012973</i>	19.78	-	38.90
<i>SNP401</i>	<i>DLAS0000287</i>	21.62	-	45.30
<i>SNP431</i>	<i>H3GA0055624</i>	22.79	-	46.30
<i>SNP437</i>	<i>H3GA0033954</i>	23.21	41.30	-
<i>SNP441</i>	<i>DLAS0004287</i>	23.83	-	48.40
<i>SNP451</i>	<i>H3GA0033984</i>	24.17	-	49.20
<i>SNP471</i>	<i>ALGA0108989</i>	25.03	-	50.10
<i>SNP561</i>	<i>H3GA0055148</i>	30.99	-	50.70
<i>SNP599</i>	<i>MARC0016326</i>	32.76	46.70	-
<i>SNP601</i>	<i>ALGA0066217</i>	33.24	-	52.70
<i>SNP611</i>	<i>MARC0084960</i>	34.15	-	53.90
<i>SNP613</i>	<i>ASGA0054380</i>	34.23	47.30	-

SNP618	ALGA0119377	35.1	48.40	-
SNP621	ALGA0111445	36	-	54.30
SNP624	DRGA0011754	36.25	49.10	-
SNP631	MARC0008973	36.76	-	56.90
SNP635	ALGA0066299	36.92	50.20	-
SNP645	MARC0091586	37.48	51.70	-
SNP651	ASGA0104264	37.87	-	58.20
SNP661	ASGA0089251	38.21	-	59.50
SNP673	ALGA0104951	38.77	55.50	-
SNP681	ALGA0123748	38.82	-	62.60
SNP691	MARC0111729	39.25	-	63.80
SNP701	ASGA0090983	39.78	-	66.00
SNP707	DLAS0004274	40.27	60.40	-
SNP711	ASGA0054736	40.51	-	68.20
SNP741	ALGA0119804	42.45	-	71.70
SNP761	ASGA0054859	43.52	-	73.50
SNP771	MARC0092941	44.59	-	74.50
SNP781	H3GA0034599	46.31	-	77.50
SNP796	ALGA0124577	47.18	72.10	-
SNP831	MARC0076283	48.55	-	79.60
SNP841	ASGA0105522	49.31	-	80.00
SNP849	ASGA0055070	50.05	74.80	-

ANEJO

<i>SNP851</i>	<i>MARC0023716</i>	50.88	-	83.10
<i>SNP861</i>	<i>ALGA0103480</i>	51.44	-	86.10
<i>SNP901</i>	<i>DLAS0003416</i>	53.32	-	87.30
<i>SNP904</i>	<i>ALGA0119023</i>	53.58	82.40	-
<i>SNP911</i>	<i>MARC0094338</i>	54.13	-	90.20
<i>SNP921</i>	<i>ASGA0099848</i>	54.5	-	90.90
<i>SNP931</i>	<i>H3GA0052996</i>	54.64	-	91.50
<i>SNP941</i>	<i>MARC0068381</i>	54.94	-	93.20
<i>SNP960</i>	<i>H3GA0035045</i>	56.18	90.40	-
<i>SNP965</i>	<i>H3GA0035082</i>	56.46	90.60	-
<i>SNP972</i>	<i>ASGA0101684</i>	56.78	91.30	-
<i>SNP974</i>	<i>ALGA0067317</i>	56.92	92.00	-
<i>SNP975</i>	<i>MARC0114025</i>	56.96	92.10	-
<i>SNP977</i>	<i>ASGA0055477</i>	57	-	95.20